


```

OY 783 CGGGCCCCGTGAGGCCCCCGGCGGTGTCGCGCGGCGTACTGAGCGCACTTCG 842
DB 1218264 CGGTCCCGCCATCCCCCGTTGCGCGCGGTGCGCCCGCATTTGTCCTGCTAGAGCCGT 1218205
OY 843 CTTCCAGCAACAACGCGGCGCGCAAGTGAAGCTGTGCGCGCAACGCAAAAGTGGCTTCG 902
DB 1218204 TAGGCGCGGCTCCGCGCGCTCGGCGGTGGCGCCCTGCGCGCGCGCGCGCGTTCG 1218145
OY 903 ACCGCGCGCGCAACGCGCGCGCGCACTGTCTTCTTCGAGCGCGCGCTGCGCGCGGGA 962
DB 1218144 ACAGCAACCCCCCGGTGCGCGCTTTCGCGCATTCGCG-----CATTTGCGCGCTTC 1218090
OY 963 GAGCCTTTCTGAGAGTGGGCGCGTCCGCGGCGTGGCGCGCGCGCGCTGCGCTTCG 1022
DB 1218089 GCGATTGCGCGCGTTCCTCCGCGCGCAACCGCGCTTGGCGCGCGCGCGCGCGCGCG 1218030
OY 1023 CATCACTGTGCGCAACGCGGCGGTCTACGCGCGCAACGAGTGCCTGCGCGCAACG 1082
DB 1218029 GAGCGCGCGGTTCGCGCGCAACCGCGCATTCGCGCAACGCGCGCGCGCGCGCG 1217970
OY 1083 GCTGC-----TCGACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCG 1137
DB 1217969 CCGCGCGCGCGGTACCAACGCGCTGCGCAACGCGCTGCGCGCGCGCGCGCGCG 1217910
OY 1138 GCGCGCGCAACGCGCTGAGCTTCAACGCTGCGCGCGCGCGCGCGCGCGCGCTGCG 1197
DB 1217909 GCGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217850
OY 1198 GGGCGTCCGCGCGCGCGCGCTGTGCGTGAACACACGAGCGAGCTGTGGCGCTTCTTC 1257
DB 1217849 GGGCGCGCGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 1217790
OY 1258 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1317
DB 1217789 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1217730
OY 1318 GCGACCAAGTCACTCATAGAGGCTCCTCAAGCGCGCTCCCAAGATGATGATTCAGATAG 1377
DB 1217729 ACCGTGAGGCGCAAGCGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217670
OY 1378 ACCTTCACTGTCAACCAAGTCTCTCTGCGATCTGAGTCAATCCCTGTGAGCGCGCGCG 1437
DB 1217669 CCGCGCGCGCGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217610
OY 1438 TCCCGCGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1497
DB 1217609 GTCCCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGAG 1217550
OY 1498 AATGCGAGTGAACGAGTGTCTTCATGAGCGAGTGAACAGCGGTATC 1545
DB 1217549 ACCCGGCGTAAAGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217502

```

RESULT 2

US-08-804-227C-7

Sequence 7, Application US/08804227C

Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Kuester, Stuart A.

APPLICANT: Rostock, Paul R., Jr.

APPLICANT: Sutton, Kimberly L.

TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: THOMAS G. PLANT 1501

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-227C-7

```

Query Match 4.8%; Score 79.8; DB 2; Length 44377;

Best Local Similarity 44.5%; Pred. No. 3.2e-06;

Matches 549; Conservative 0; Mismatches 667; Indels 19; Gaps 5;

```

OY 21 GACCTTGCAGAACCCGAGCCCAACGCGCGCGCTCTTGCACACCGCGCGTGTGGCGCC 80
DB 23286 GCGCGTGCCTGACCGCGCACCTTGCCTGCTCAAGCGCTTGCAGACTCGGCGAGC 23345
OY 81 GCGCGCGAGGAGCGCGCGCTGTGGGAGAGCGCGCGCTTTCACCGCGCGAGCGCAAG 140
DB 23346 GCGACCGCTGTGACCTGTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 23405
OY 141 CAAGAACGTGCGCGCTGACCGCGCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 200
DB 23406 CACCGCGCGTGCAGCGTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 23465
OY 201 TGGCGTCACTTCAACGAGCGCGCGCGCGCGCTGTAAGAGAGTGTGCGCTGTGCT 260
DB 23466 GTGGGCGCGCGCTTCACTTGCCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 23525
OY 261 GCGCGTGTGCGCGCTGTGAGCGCGCGCGCGCGCGCTTTCACCGCGCGAGTCCGTC 320
DB 23526 CCGGTCTTGTGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 23585
OY 321 GCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 380
DB 23586 CCGCGCGCGCGGTCTGTGCGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 23645
OY 381 CTACTGGCGCAAGGCACTGCGCGAGAACCTGTGCGCGCGCGCGCGCGCGCGCGCTTACTG 440
DB 23646 CACCGTCTTCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 23705
OY 441 GCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCA 500
DB 23706 CGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 23765

```

OY	50	CTGGGCGGTGAGCGCGTGGGCGCGCGCGCTCTGGGCGCTCATTTGATGATCTAAGGCGATCAACGA	560
Db	23766	CGCGGAGCTAAGCCGCGCTCTCTCTGTCGCCCGTGGCGCTGCGAGTCAACCGACCGCGCGC	23825
OY	561	CGAGTGCAGACTTCTTGAGAGACGCGCTTTCGCTGACAGCTGACGCGCCGCGCGCTGACGCA	620
Db	23826	ACTGGCGCGCGCTGCTGCA-----CGAGCAGCGCGCCACCGTGGCGGTGACCGCC	23874
OY	621	GGCCCCGTTTACGGCGCTTGCCTTGCACCGCCAGACGCAACGCGGCGCAATTGACACAA	680
Db	23875	CCGCGCTCTGGTGCCTCCCTGGGCGCGCGCTGGCGGAGAACGCGACCGCG--CGAATGCGCGCGCG	23933
OY	681	CGAGCTGAGAGAACAAACGAGTGGTGGCGCAAGCTGGGGCGCATGGGCGCTGGGCGCGCGCC	740
Db	23934	CTTGGCGCGCAAGACACAGCGCGCGCGCGCGCATCTGGTGCACCTGGACCGCGCGCGCTT	23993
OY	741	GGGCGCACCGCGCAGCGCCAGCGCGCGCGCGCGCGCATTTCTGGTGGCGCGCCCGGTGAGCGGCC	800
Db	23994	CGAGCGCGCTGGTGTGCTTCTCTTGGGTCTCGGAGTGGGGGCGCGCGCGCCAGGGGCGG	24053
OY	801	GCGGCGCGCGCTGCTGCGCGCGCGCTACTGAGAGCGCGACCTGGCGTTTCAACGCAACGCGG	860
Db	24054	CTAAGCGGCGCGCGCACGCGCGCGCATCTGACGCGCTGGCGCGAACGCGCGCGCGCGCGCGGGGT	24113
OY	861	GGCGGAGTGAAGCGCTGTCCGCGCGCAACGCGCAAAATGGGCTGGCGACGCGCGCGCGCGAGGCGG	920
Db	24114	GGCGCGGTTCTCCGT---GGCTTGGAGACCGCTTGGGCGCGAGGACGCGCGCGCGAGGTGC	24170
OY	921	CCGACGACCTGGTCTTCTCTCGAGCGCGCCGCTTGGCGCGCGCGCGCGAGAGCGCTTCTGTGAGAGT	980
Db	24171	CGAGCGCGAGATCTCTTACGCGCGGCGCGGCGCTGGCTCCCTCTGACCGCCGACAGGCGGTGGG	24230
OY	981	GGGCGCTTCGAGGCGTGGCGCGCGCGCGCGCGCTGGCGCTTGGCATTCAGTCTGGCGACCC	1040
Db	24231	GACCGTTCGCGCGCATGCTCTGGAGGCGCGCGCACGCGCTGCGTGGTGGCGAGTGGAGTGG	24290
OY	1041	GGGCGTCTACGCGCGCGCAACGAGCTGGCCCCCGCGACCGAGACGCGCGTGGCTGACCGGAAAGA	1100
Db	24291	GAGCGCGGTTCGCGCGCTCTCTACACCTGGGAGCGCTCCCGCGTATCTTTGAGAGCAATCCC	24350
OY	1101	GTACTGGGTGGTGGCGCGC---GCCGCGCGCCGTGGCGAGCGCGCGCGAGCGCGCTGAGCTT	1157
Db	24351	GGAAGTGCAGGCGGTGGCGCGCGCGCGCGCAACTCGCGCGCGAGACCGGAGACTTCGACCACTTC	24410
OY	1158	CAGCGTGCAGCGCGCGCGCGCGAGCTGC--TCTGGGCAATCAACGCGCGCTTCGCGCGCGCGCC	1216
Db	24411	CGAAGCTGTCGCGCGAGCTGACCGGCGAGTCCGGCGCAAGACGCGACGCAACTCTGTGG	24470
OY	1217	TGCTGTGCGCTGACACCAACGACGCGCTTCTGGGCG	1251
Db	24471	GCTGGTGCAGCAACGCGCGCGCGCTCTCTGAGC	24505

RESULT 3
 US-08-804-198-1
 : Sequence 1, Application US/08804198
 : Patent No. 5845320
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Burgett, Stanley G.
 : APPLICANT: Kutsche, Stuart A.
 : APPLICANT: Rao, Nagendra R.
 : APPLICANT: Richardson, Mark A.
 : APPLICANT: Roseck, Paul R., Jr
 :
 : TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE
 :
 : NUMBER OF SEQUENCES: 6
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSSEE: PAUL R. CANTRELL 1138
 : STREET: LILLY CORPORATE CENTER
 : CITY: INDIANAPOLIS
 :
 : STATE: IN
 :
 : COUNTRY: USA
 :
 : ZIP: 46285
 :

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 COMPUTER: Macintosh
4 OPERATING SYSTEM: Macintosh 7.0
5 SOFTWARE: Microsoft Word 5.1
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/804,198
8 FILING DATE:
9 CLASSIFICATION: 435
10 ATTORNEY/AGENT INFORMATION:
11 NAME: CANTRELL, PAUL R.
12 REGISTRATION NUMBER: 36,470
13 REFERENCE/DOCKET NUMBER: P9113
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 317-276-3885
16 INFORMATION FOR SEQ ID NO: 1:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 44377 base pairs
19 TYPE: nucleic acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 MOLECULE TYPE: DNA (genomic)
23 FEATURE:
24 NAME/KEY: CDS
25 LOCATION: 350..14002
26 FEATURE:
27 NAME/KEY: CDS
28 LOCATION: 14046..20036
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: 20110..31284
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: 31329..36071
35 FEATURE:
36 NAME/KEY: CDS
37 LOCATION: 36155..41830
38
39 US-08-804-198-1

```

```

Query Match      4.8%: Score 79.8; DB 2; Length 4437;
Best Local Similarity 44.5%; Pred. No. 3.2e-06;
Match 549; Conservative 0; Mismatch 667; Indels 19; Gaps 5;

Dy      21  GACCTTGCCAGACCCGAGCCACCGAGCGCGCTCTTGAGCCACCAGCGCGTGTCTGCGGCC 80
Db      23286 GCGGTGTGCGGTGACCGCGACCCGACCCGTGGCTCTGTCCAGGCTTGAGCGACCTCGGCAACAC 23344

Dy      81  CGGCCCCGAGCGACGCCCCGGTCTTGAGCGAGGCGCGCGCTTCCAGCGCGACGACCAAG 140
Db      23346  CGCACCCCTGTGAGACCTTCACCTTGCGGCGCGCGGTGCGACACGCCCCCGACCAACATGCGGT 23400

Dy      141  CAAGAACGTGCGGTGAGCGGCGCATCTGCGCGCGCGGCGCACACGAGCGCAACAGCTTCTGCA 200
Db      23406  CACGCGCGGTGCCAGCTGTAGGGGCTGAGGCGGGGTGAGCGCGCTTGAGACTGTGCCAGGT 23460

Dy      201  TGGCGTCACTGTACAGGAGCGGCCCATCCGCTGTACAGAGAGGTGCGGCTGCGCTGTGT 260
Db      23466  GTGGGGGGCGGCTCATTCAGACTTCCCGGCGGCGCGGAGCGCCGGGTCTTGAGCGGTCTGCG 23520

Dy      261  GCGGTGTGGCTCTTGAGTGAAGCGGCGCGCTGCGCTTGGGCTTACCGCGCAAGATCCGTCT 320
Db      23526  CGGGGTCTCTCCCAACCCCGCGCGGAGGACAGATGCGCGTACGAAATGGCGGCGGTCTT 23580

Dy      321  GCTCATGAGCGCCGAGGACATCCCAAGTACGCTGCGCCGGAACCTGTGTCAACGCGGCGG 380
Db      23586  CGGCGCGCGGGTCTTGAGGAGAACCGGCGCGACTCCCGGCTCCCGGCGCTGAGCGCGCGCGG 23640

Dy      381  CTACTGGGCGCAAGGCACTGCCCGAGAACCTTGCGGCTGCGCGGACAGGATGTGCGCTTACTG 440
Db      23646  CACGTCCTTATGCGCGGCGACCTCAAGAGGATCCCGGCGGACTGTGATCGGATCTCTCT 23700

Dy      441  GCGGACCGCGACGCGCGCGGTCTTCTACAGGTGAAGAGCGGCGAGCGCGGATGCTCTTCCA 500

```

Db	23706	CGAGGACCGGGCGGAGCCGGGTGGTGGCCGGACCCGACGCCGCCCGACAGCGCGCG	23765
QY	501	CTGGCGCGTGGCGCGTGGAGCGGCCCTCTGGGCGCTCAATTGATGTCATGCAATCCGA	560
Db	23766	CGCGCGACCTAACCGGGCGTCTCCCTCGTCCCGCGCCCTCGACGTCAACGACCGCGCG	23825
QY	561	CGAGGTACAGCTTCTGGAGAGCGCGCTTTCGCTGACAGCTGAGCGCCCGCGCGCTGAGCA	620
Db	23826	ACTGGCGCGCGCTGTGCA-----CGAGCAGCGCGCCACCGTCCGCGGTGACCGCC	23874
QY	621	GCGCCGCTTGAGCGCGCTGCGCTGCGCGCCAGCAGCCACGACGCGGCCAACTTGACAAACA	680
Db	23875	CCGCGCCTGTGGTCCCGCTGGCGCGCGCTGCGGGAGACGGACACCGG--CGACATGCGCGCG	23933
QY	681	CGAGCTCGAAGAACACCAAGTGGTGGCCAAAGCTGGGCGCACCTTGGGCGTGGGCGCGCCC	740
Db	23934	CTTGGCGCGCCAAAGACCAAGCGCGCGCGCGCCACCTGGTGCACCTTGGCGCGCGCGGCT	23993
QY	741	GGGCGCCACCGCCAGACCGACCGCGCGCGCGCGCATTCGTGGCGGGCCCGGTGAGCGCC	800
Db	23994	CGAGCGCGTGGTGGTGTCTTCTCTCGGTCTTCGGAGTGTGGGCGCGCGGCCCAAGGCGG	24053
QY	801	GCGGCGCGCGTGTGTGCGCGCGCGCTGACTGAGAGCGCACTTGGCTTCAAGCAACGCGG	860
Db	24054	CTAACGCGGCGCGCAGCGCGCGACCTCGACGCGCGTGGCGGAACGCGCGCGCGCGCGGGGT	24113
QY	861	GCCCGACGTAGAGCTGTGCGCGCCAGCGCAAAATGGTGTGCGCACCGCGGCCCGGACGCGG	920
Db	24114	GCCCGACCTTCTCGGT---GGCTCTGAGAGCCCTTGGGCGCGGACCGCCCGCGCGAGTGC	24170
QY	921	CCGCAACGCTGTCTTCTCCGAGCGCGCGCTGCGCGGCCCGGCGAGAGGCTTTCGTGGAAGT	980
Db	24171	CGAGGCGGAATTCCTCAGCGCGCGCGCGGGCGTGGTCCCTCTCGACCCCGACAGCGGTGCG	24230
QY	981	GGGCGGTCCGAGGCGTGGCGCGCGCGCGCGCGCTTTCGAGATCAGTGTGACCC	1040
Db	24231	GACCTTCGCGCGCATGTGTGAGAGCGGGCAGCGCGCTGGTGGAGTGCAGATGAGTG	24290
QY	1041	GGGGTGTCTAACGCGCCACAGAGCTGCGCCCGCCAGACCGAGACGCGTGTGACACGGAAAGA	1100
Db	24291	GAGCGGGTGTGCGCGCTCTCAACCTTGGAGTGTCCCGCGCTGATCTTTGAGAGCAATCCC	24350
QY	1101	GTACTGGTGTGTGGCGCGC---GCGGGGCCCTGTGCGAGAGCGGCGGCGACGCGCTAGCTT	1157
Db	24351	GGAAGTGCAGCGGTGCGCGCGCGCGCGCGCACTTCGCCCGAGACACCGAGATCTGACACATTC	24410
QY	1158	CAGCGTGGCGCGCGCGCGCGAGCTGC--TCTGGGCAATCAACGCGCGTCCGCGCGCGCGC	1216
Db	24411	CGAACTGTCTCGCGAGCTGACCGCGCAGTCCGGCGCAAGACGCGACGCGACCTGTGTGG	24470
QY	1217	TGCTGTGCGTGCACACCAAGCGAGCGGTCTGGGCGC	1251
Db	24471	GCTGTGTGCGCGACACGCGCGCGCGCGCTCTCGGAC	24505

```

RESULT 4
US-09-105-537-34
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
/

```

! ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match	4.7%	Score 79.4	DB 3	Length 4689
Beet Local Similarity	44.1%	Pred. No. 3.5e-06		
Matches 427	Conservative	Of	Mismatches 536	Indels 6
				Gaps 2
QY	317	CGTGCCTCATGAGCGCCCAAGGACATCCCCAAATGACGCTGCGCCGAGCCTGTGTCACGCGGC	376	
DB	2723	CGCCACACAGCGCGCCGAGACGATCGTGGGTATACCGCGTCGATGAGGACCGCTGACCGGCT	2782	
QY	377	CGGCTACTGAGGCCAAGCACTGCGCCGAGAACCTGCGGCTGCGGACACAGTGTCTGGCT	436	
DB	2783	CGGCGACAGCGGACCTGTCTCGGAGGGTGGAAATCGTCGCGCTCGAGAGGAGGACAGAGCCG	2842	
QY	437	ACTGAGGCGAACCAGCGACGCGCGCGTGTCTTACAGCGTGAACGACGCGGAGCGGCTGTCT	496	
DB	2843	AGCTGTGTGGCGCGCTGAAGGCGCGGAGAGCGGAGGTGACGTACTGTGAAGCCGCGCGCG	2902	
QY	497	TCGACTGCGGCGTGGCGCGTGGAGCGACCGCGCTCTGAGGCGCTCATATATGTCTTACGACATCA	556	
DB	2903	ACGACGACCGTGAAGGCGCTCGCGCGCGGCTTACCGGACTGACGACGCGGACCGGCTTCA	2962	
QY	557	CGGACGAGGTGACACTTCTGTGAAGCGCCTTTCGTCGACGCTGACGCCCGCGCGCTTCA	616	
DB	2963	CGGCGGTGTCTCGCTCTCTGACGACCTCGTGCCACAGTTCGCTGTGGTGCACGACACTCG	3022	
QY	617	GCACAGCGCGCTTTCAGCGCGCTGTCGCGCGCCAGACGACGACGACGCGGACCACTTTCACA	676	
DB	3023	GCAGCGCGGAAATGAAGCGCGCGCGCTGTGTGTTCGTGACCCAGGCGCGCGTCTCGCTCGGAC	3082	
QY	677	ACAAAGAGCTTGAGAACACCAAGTGTGTGCACAGCTGAGCGACCTTGGCGCTGGCGCGCG	736	
DB	3083	GTCTTCGACACCCCGCGGACCCCGACCGGCGCAAGCTCTGGGCGCTGGCGCGCGCTGTGCG	3142	
QY	737	CCCGCGGCGCACCGCGCACGCGGACGCGCGCGCGGCGCACTTCGCTGGCGGCGCGCGT-AG	795	
DB	3143	CCCTTGAAGCACCCGMAAGCTGGGCGCGGCTGTGTGACCTCCCGCGCACGCGCGCATGGCG	3202	
QY	796	CGCCCGCGCGCGCGCGTGTGTGCGCGGCGCTACTTGAAGCGGACCTGCGCTTTCACGACACA	855	
DB	3203	CGCGCGTTCGCCCACTGTGTGACCGGACCTTCGCGGCGCACCGGCGAGGACAGATTCGCCA	3262	
QY	856	CGCGGCGCCCAAGTGAAGCTGTCCGCG-----GACCGCGAAGTGGCGCTGCGCACCGCGCG	910	
DB	3263	TCGCGACACAACGGAATCTCAGCGCGCGCGCTGCGCGGACCGCTTCCACGAGGTGCGG	3322	
QY	911	CGGACGCGCGCGGACGCTGTGTCTTCTCCGAGCGCGCGCTGCGCGCGCGCGGACGACCTCT	970	
DB	3323	CCACCGCGGACGTGGGACGCCCAACGCGACGCTCTCATACCGCGCGGACCGGACCGCTTCG	3382	
QY	971	TCGTGAGAGTGGGCGCTTCGCGGAGCTGGCGGCGCGCGCGCGCGCTGGCTTTCGACATCACT	1030	
DB	3383	GCAACGACGCGGACGCTGGAATGCGCCACACAGGAGCGMAACACTCTCTCTCGTCAAGCC	3442	
QY	1031	CGTGCAGCCCGGCGGTGTACGCGCGCAACGAGCTGCGCGCGCGGACCGGACGCGCTGCG	1090	
DB	3443	GCAGCGGCGAACAAGCCCGCGGAGCACCTCATCGCGGAACTCACGCGCATTCGAGCG	3502	
QY	1091	ACGCGAAGATGACTGTGGTGTGTGCGCGCGCGCGGCGCGGCGCGGACGCGCGGACGCGC	1150	
DB	3503	CCCGGTGACATGCGCGCGCTGTGCGACGCTGCGACCCCGACGCGCATGCGACCTCTCTCG	3562	
QY	1151	TCAGCTTACGCTGCGGCGCGGCGGACGCTGCTCTGTGGGATACAGGCGGTTCGCGCG	1210	
DB	3563	ACGCGATTCGCCGCGGAGACGCGCTTACCGCGCTGTGTACACGCGCGGCGCACCTGGCG	3622	
QY	1211	GCAGCTGTGTGCTGACACACAGAGGCGCTCTGGGCGCTTCTTTCGCGCTGCGCGCGCG	1270	
DB	3623	GCATTCGCGTGAAGCTCACCGCGCGCGGAGACATTCGCCGCACTCTGGGCGCGGAAAGCA	3682	
QY	1271	GCCTTCGCGG 1279		

Db 3683 GCGGCGCCG 3691

RESULT 5

US-09-105-537-5

Sequence 5, Application US/09105537A

Patent No. 6265202

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600.438U51

CURRENT APPLICATION NUMBER: US/09/105,537A

CURRENT FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 5

LENGTH: 36778

TYPE: DNA

ORGANISM: Streptomyces venezuelae

US-09-105-537-5

Query Match 4.7%; Score 79.4; DB 3; Length 36778;
Best Local Similarity 44.1%; Pred. No. 3.7e-06;
Matches 427; Conservative 0; Mismatches 536; Indels 6; Gaps 2;

QY 317 CGTGCCTATGAGCCGCCAGGACATCCCAATAGCCCTGCCGCCGACCTGTCTACGCGGC 376
DB 29713 GCGCCACAGCGCCGCCAGCACTGAGCGGTACCGGTGAGTGAAGCCGCTGAGCGCCT 29772
QY 377 CGGCTACTGAGCCCAAGGCACTGCGCGAAGACCTGCGCGTGCACACAGGTGCTGCGCT 436
DB 29773 CGGCGAAGCGGACCTGTCTCGGCGCGGTGATGTGCGCGTGCAGGACCAAGAACCG 29832
QY 437 ACTGGGCGGACCGGCGCGCGGTGTCTACAGCGTGAACGACGCGGACCGGTGCTCT 496
DB 29833 AGCTCTGCGCGCGCTGAAAGCGCGGAGCGAGGTGACGTACTGAAAGCGGCGCG 29892
QY 497 TCCTATGCGCGGTGCGCGGTGCGCGCGCGCTGTGCGCGCTATGATGTCTACGCGATCA 556
DB 29893 ACGAGACCGGTGAGGCGCGCGCGCGCGCTGACCGCACTGACGACCGCGCGCTTCA 29952
QY 557 CGGACGAGGTGACGCTTCTGAGAGCGCGCTTGTGACACGCTGACGCGCGCGCGCTCA 616
DB 29953 CGGCGGTGTCTCGCTCTCTGACGACCTGTGACAGGTGCGCTGCGGTGACGCGACCTG 30012
QY 617 GCGAGCGCGCTTCAAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
DB 30013 GCGAGCGCGGATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30072
QY 677 ACAACGAGCTGAGAAACAACGAGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 736
DB 30073 GTCTGACACCG 30132
QY 737 CCG 795
DB 30133 CCGTTGAGACCG 30192
QY 796 CG 855
DB 30193 CG 30252
QY 856 CG 910
DB 30253 TCAGGACCG 30312
QY 911 CGGACG 970
DB 30313 CCAACCG 30372
QY 971 TCGTGAAGTGGCG 1030

Db 30373 GCGGCGCCG 3691
QY 1031 CGTGCCTATGAGCCGCCAGGACATCCCAATAGCCCTGCCGCCGACCTGTCTACGCGGC 1090
DB 30433 GCGGCGCCG 3691
QY 1091 ACCG 1150
DB 30493 CCG 30552
QY 1151 TCAGCTTACG 1210
DB 30553 ACG 30612
QY 1211 GCGGCG 1270
DB 30613 GCGGCG 30672
QY 1271 GCGGCGCGG 1279
DB 30673 GCGGCGCGG 30681

RESULT 6

US-09-320-878-19

Sequence 19, Application US/09320878A

Patent No. 6117659

GENERAL INFORMATION:

APPLICANT: ASHLEY, Gary

APPLICANT: BETLACH, Melanie C.

APPLICANT: BETLACH, Mary C.

APPLICANT: MC DANIEL, Robert

APPLICANT: TANG, Li

TITLE OF INVENTION: RECOMBINANT NARHONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 30062202120

CURRENT APPLICATION NUMBER: US/09/320,878A

CURRENT FILING DATE: 1999-05-27

EARLIER APPLICATION NUMBER: CIP OF 09/141,908

EARLIER FILING DATE: 1998-08-28

EARLIER APPLICATION NUMBER: CIP OF 09/073,538

EARLIER FILING DATE: 1998-05-06

EARLIER APPLICATION NUMBER: CIP OF 08/846,247

EARLIER FILING DATE: 1997-04-30

EARLIER APPLICATION NUMBER: 60/119,139

EARLIER FILING DATE: 1999-02-08

EARLIER APPLICATION NUMBER: 60/100,880

EARLIER FILING DATE: 1998-09-22

EARLIER APPLICATION NUMBER: 60/087,080

EARLIER FILING DATE: 1998-05-28

NUMBER OF SEQ ID NOS: 34

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 19

LENGTH: 38506

TYPE: DNA

ORGANISM: Streptomyces venezuelae

US-09-320-878-19

Query Match 4.7%; Score 79.4; DB 3; Length 38506;
Best Local Similarity 44.1%; Pred. No. 3.7e-06;
Matches 427; Conservative 0; Mismatches 536; Indels 6; Gaps 2;

QY 317 CGTGCCTATGAGCCGCCAGGACATCCCAATAGCCCTGCCGCCGACCTGTCTACGCGGC 376
DB 27855 GCGGCGCCG 3691
QY 377 CGGCTACTGAGCCCAAGGCACTGCGCGAAGACCTGCGCGTGCACACAGGTGCTGCGCT 436
DB 27915 CGGCGAAGCGGACCTGTCTCGGCGCGGTGATGTGCGCGGTGAGGCGGCGGCGGCG 27974
QY 437 ACTGGGCGGACCGGCGCGCGCGGTGTCTACAGGTGAACGACGCGCGCGGTGCTCT 496
DB 27975 AGCTCTGCGCGCGCTGAAAGCGCGGAGCGAGGTGACGTACTGAAAGCGGCGCGG 28034

Db	28635	CCGCGTCACACATGCGCGCCTTGCGACGTGGCCGACCCCAAGGCATGCGCACCTTCTTG	28695
Qy	1151	TCAGCTTCACGCGTGCAGGCCCGGCGGCGACGTGCTCTTGGCATCAACGAGCGTCCGCGG	1210
Db	28695	ACGCGATCTCCCGCGGAGACGCCCTTCAACGCGCTGTCTCACACCGCGGCGGACCGGGG	28755
Qy	1211	GCCGCGTGTGTGCGGTGAGACCAACGACGAGCGCTCTGGGCTTCTTTCGCGGTCGCGGCG	1270
Db	28755	GCGATCCGCGTGGACGTCAACGCGGCCGAGGACATGCGCGCATCTTGAGGCGCAAGACA	28815
Qy	1271	GCGTTCGCGG	1279
Db	28815	GCGGCGCG	28823

```

RESULT 8
US-09-657-440-19
; Sequence 19, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062200210
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-19

```

Query Match	Similarity	4.7%	Score 79.4	DB 4	Length 38506
Best Local Match	Similarity	44.1%	Pred. No. 3.7e-06		
Matches 427	Conservative	0	Mismatches 536	Indels 6	Gaps 2
Qy	317	CGTGCCTCATGAGCGCCCGACGACATCTCCCAAGTACGCGCTGCCGACCTTGATCAAGCGGC	376		
Db	27855	CGCCACACAGCGCGCCGACGACATCTGAGGCTTACCGCGTCTGAGTGAAGCCGCTACCGGCTT	27914		
Qy	377	CGGGCTACTGGGCGCAAGGACTGCGCCGAGAACCTTGGGCGCTGCGGACACAGGTCTGGCTT	436		
Db	27915	CGGGCCAGCGCGACCTTCTCCGGGGGTTGGAGTCTGCGCGCTGGAGAGGACCGAAGACCG	27974		
Qy	437	ACTGGGCGGACCGGCGGACCGCGCGCTGTTTCTACAGCGTGAACGACCGGCGACGCGGTCTCT	496		
Db	27975	AGCTGCTGAGGGCGGCGCTGAAGGCGCGCGGAGCGAGGTCTGACGTACTGGAAGCGGGGCGG	28034		
Qy	497	TCCACTGGCGGCGTGGCGCGGTGGGCGGCGCGCTCTTGGGCGCTTCATGATGTCTACGGCATCA	556		
Db	28035	ACGACGACCGGTGAAGGCGCTTCCGCCCGCGGCTCTACCGCATGTACGACCGGCGGAGCGGCTTCA	28094		
Qy	557	CCGACGAGAGTTCACGCTTCTTGGAGAGCGGCTTCCGTGACACGCTGACCGCGCGCGGCTCTCA	616		
Db	28095	CCGGCGTGTGTCTCGCTCTCTGACGACCTCTGTGACACAGTGTGCTGTGGTGAAGGCACTCG	28155		
Qy	617	GCGAGGCGCGCTTTCAGCGGCTGCTGCGCGCCGACGACGACGACGCGCGGCAACTTTCGACA	676		
Db	28155	GCGACGCGCGGATCAAGGCGCGCGCTGTGTGTCCGTCACCGAGGGGCGGATCTCCGTCCGAC	28214		
Qy	677	ACAAAGAGCTTGAGAACACCAAGTGTGGTGGCGACAGTGGGCGCACTTGGCGCTGGGCGCGG	736		
Db	28215	GCTCTGACACCCCGCGGACCCCGACCGGGCGCAAGTCTTGGGGCTTGGCGGTGGCTGG	28274		

OY	73	CCCCGGGCCCCAACGCGACGGCGAAGCGCGCGGCGCGCGCATTCCTGTGCGGGGCCCCCGG-AG	795
Db	28275	CCCTTAGGCAACCCCGAAGCTGGAGCGGGCGCTGTCACTCCCGGCCACGCCCGATGCGCG	28334
OY	796	CGCCCCGCGCGCGCGGTGTGTCCGCGCGCGCTACTGGAGGCGGACCTTCGCTTTCACCGCA	855
Db	28335	CCGCGCTTCGCGCGACCTGTCACTCCCGCGCGCGACCGCGGAGGACAGATGCGCA	28394
OY	856	CGCGGGCCCCGACGTGAAGCTGTGCGC----GACCGCGAAAGTGGCGTTCGCGACCGCGGC	910
Db	28395	TCCGCAACCAACGGACTTCACAGCGCGCGCGCTCGCCCGCGCACCCCTTCACGAGACTGCGCG	28454
OY	911	CCGACCGGCGCGCGCGACGCTGGGTCTTTCGCGAGGCGCGCGCGCGCGCGCGCGGACGAGCGCTT	970
Db	28455	CGACCCCGGACTGGACAGCCCCCAAGGCAACGCTCTTCAACCGCGGCGACCGGAGCCCTTG	28514
OY	971	TCTGTGAGGTGGGCGGTCCCGGGGCTGGCGGCGCGCGCGCGCTTGCGCTTTCGCGATCACTGT	1030
Db	28515	GCAAGCGACCGCGCGCACGCTGGATGAGGCCACAGAGAGCGGAAACACTCTCTCTGTATAGCC	28574
OY	1031	CGTGGCAACCGGGCGGTGTACCGGCGCAAGAGTGTCCCGCGGACCGACGCGCTGCG	1090
Db	28575	GCAAGCGCGGAAACAAGCCCCCGGAGGCGACCACTCACCGCGGAACTCACCGGATCGGCGG	28634
OY	1091	ACCGCAAAAGATACTGGGTGTGTGGCGCGCGCGCGCGCGCGCGCGGAGACCGCGGACCGCG	1150
Db	28635	CCGCGCGTACCATTCGCGCGCGCTGGAGGTGCGGACCCCGACGCGACATGCGGACCTCTTCG	28694
OY	1151	TCAAGCTTCAGCGCTGGCGCGCGCGGCGGAGCGTGTCTCTGGGACATCAACGCGGCGTCCGCGCG	1210
Db	28695	ACGCGCACTCCCGCGCGGAGACGCGCCCTTCAACCGCGCGTGTCCACACACCGCGCGCGCACCGGCG	28754
OY	1211	GCGGCGCTGTGTGTGTGACACCAACGAGGCGCTTGGGCGTTTCTTCGCGCGTGGCGCGCGG	1270
Db	28755	GCGATCCGCTGTGACGTCAACGCGCGCGGAGGACATCCGCGCATCTCGGCGCGGAGACGA	28814
OY	1271	GCGTCCGCGG 1279	
Db	28815	GCGGCGCGCG 28823	

RESULT 9
 US-08-403-852D-6
 Sequence 6, Application US/08403852D
 Patent No. 5891695
 GENERAL INFORMATION:
 APPLICANT: Blanc, Veronique
 APPLICANT: Bianche, Francis
 APPLICANT: Crozet, Joel
 APPLICANT: Jacques, Nathalie
 APPLICANT: Lacroix, Patricia
 APPLICANT: Thibaut, Denis
 APPLICANT: Zagorec, Monique
 APPLICANT: Debussche, Laurent
 APPLICANT: De Crecy-Lagard, Valerie
 TITLE OF INVENTION: Polypeptides Of Streptogramins
 TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
 TITLES OF INVENTION: Coding For These Polypeptides And Their Use
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Finnegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/403,852D
 FILING DATE: 10-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR 93/00923
 FILING DATE: 25-SEP-1993
 PRIOR APPLICATION DATA: FR 92/11441
 APPLICATION NUMBER: 25-SEP-1992
 FILING DATE: 25-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03806.0054-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1833 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: S.priestinaespiralis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 103..1689
 US-08-403-852D-6

Query Match 4.7%; Score 78.6; DB 2; Length 1833;
 Best Local Similarity 45.4%; Pred. No. 4.8e-06;
 Matches 399; Conservative 0; Mismatches 474; Indels 6; Gaps 3;

414 GCTGCGGACACGCGTGTGCGCTACTGAGCGCCGACCGCCGCGTGTCTTCAACGCGT 473
 408 GCTGCTGATGCTGCGACTGCTGCTCTTCCGCGCGCGCTGCGCTGCGCGCTTCCGCCA 467
 474 GAACGACGCGGACGCGGTGCTCTTCCACTGCGCGGTGCGCGCGCGCGCGCTCTGAGG 533
 468 GGACTCGCGCACTATCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 534 GCTCATTTGATGTCTACCGCATCAACGACGAGTGCAGCTTCTGAGAGCGCTTCTGCTGA 593
 528 GACCAACACCGCTGCGCGCTCATGACGAGTCTTCAACGACGAGAGCGCCCGCGCGCAT 587
 594 CAGCGTGAACGCGCGCGCGCTCAACGACGCGCGCGCTTCAACGAGCGCTTCCGCGCAG 653
 588 CGGCTGTGAGGAGCGCGCGCTCACTGAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCG 647
 654 CCAAGACGCGCGCACTTGAACAACGAGCTTCAAGAACAGAGTGTGAGCGCAAGT 713
 648 CTTCTGTGACCACTTCTGTGAGGCTCATCTTCTGATCACTGCGCGCTGCGCGCTGCT 707
 714 GGGCACTGCGCGTGGCG 773
 708 GGGCTGTGCGCGTGG - CCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 765
 774 CATTCTGTGCGCGCGCGCGTGAAGCGCGCGCGCGCGCGCGCGCTGCTCTGAGAGCG 833
 766 CCGCACTGTGCGCGCGCGCTGCTTCAACCTCGGCACTGAGCGCGCTGCTGAGCGCATC 825
 834 CAGCTGCGCTTCAAGCAACAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 893
 826 ATCTCTGCG 884
 894 GAGCTGCGCAACG 953
 885 GAGCG 944
 954 GCCCGGGAAGCTCTTCTGTGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1013

DB 945 CATGAGCTTCTTCAACCGACCGCGCTTCAACGAGCGCGCTGCGCGCGCGCGCGCGT 1004
 QY 1014 GGCCTTCCGCACTACGTGTGAGACCGCGCGCGTGTACGCGCGCAAGTGTGCGCGCGA 1073
 DB 1005 CGGCAATGCGCGCTGCTCTTCTGCTCAACCGACGACCTCCAACTGTCTGCGCTACGA 1064
 QY 1074 CCAACGCGCTGCTGACCGCGCAAGAGTACTGTGTGTGCGCGCGCGCGCGCGCGTGC 1133
 DB 1065 CCGCTTGAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1124
 QY 1134 GAGCG 1190
 DB 1125 GCGCGCGCTTCCG 1184
 QY 1191 CATCAACGCGCGCTGCG 1250
 DB 1185 CATGACACTGTGCG 1244
 QY 1251 CTCTTCCG 1289
 DB 1245 CGGCTACG 1283

RESULT 10
 US-08-510-646B-6
 Sequence 6, Application US/08510646B
 Patent No. 6077699
 GENERAL INFORMATION:
 APPLICANT: Blanc, Veronique
 APPLICANT: Crouzet, Joel
 APPLICANT: Jacques, Nathalie
 APPLICANT: Lacroix, Patricia
 APPLICANT: Thibaut, Denis
 APPLICANT: Zagorec, Monique
 APPLICANT: Debussche, Laurent
 APPLICANT: De Crecy-Lagard, Valerie
 TITLE OF INVENTION: Polypeptides Involved In The
 TITLE OF INVENTION: Biosynthesis Of Streptogramin, Nucleotide Sequences
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESSES:
 ADDRESSER: Pimegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/510,646B
 FILING DATE: 03-AUG-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/403,852
 FILING DATE: 10-MAY-1995
 PRIOR APPLICATION DATA: PCT/FR 93/00923
 APPLICATION NUMBER: 25-SEP-1993
 FILING DATE: 25-SEP-1993
 PRIOR APPLICATION DATA: FR 92/11441
 APPLICATION NUMBER: 25-SEP-1992
 FILING DATE: 25-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03806.0054-01000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4400
 TELEFAX: (202) 408-4400

GENERAL INFORMATION:
APPLICANT: Leopardi, Rosario
TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,473
FILING DATE: 26-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARCD:239
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-690-473-1

Query Match 4.6%; Score 77.6; DB 2; Length 4257;
Beet Local Similarity 43.3%; Pred. No. 7.4e-06;
Matches 548; Conservative 0; Mismatches 694; Indels 24; Gaps 3;

QY 8 ACACGGTGCACCGAGCCCTGCGACGACCGGAGCCGAGCGGCGGCGCTCTGCGACCGCGAC 67
DB 1853 ACAGAGGGGTGCGCCCGCTCGCGCGCGACACCGGAGAGCGCGGTGCGCGCGGT 1912
QY 68 CGTGTGCGGCG 127
DB 1913 ACGGGCGCGCGGAGATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
QY 128 CGCAGGCGCAAGGCAAGACGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
DB 1973 CGGGGGGCGACGACCG 2032
QY 188 ACAGCTTGTGCAATGCGCTTCACTTCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 247
DB 2033 GCGCGCGAGCG 2092
QY 248 GCGTGCCTGCTGCG 307
DB 2093 CGCTGCGCGAGGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2152
QY 308 CGCAGCATCGCTCGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
DB 2153 CGCGCAGCG 2212
QY 368 TCACG 427
DB 2213 CGACG 2269
QY 428 TCGTGCCTTACTGCG 487
DB 2270 TGGTGTCTATGCG 2329

QY 488 CGGTGCTCTTCACTGCG 547
DB 2330 CGCGCGTGCAGCG 2389
QY 548 ACAGCAT-----CACGACGAGGTGCAAGCTTCTGAGAGCGCGCTTTCG 589
DB 2390 CGCGCTGCGAGCTCG 2449
QY 590 CTGACAGCTGAGCG 649
DB 2450 TCG 2509
QY 650 GCACGCAAGACG 709
DB 2510 CGCGCGCTCG 2569
QY 710 AGCTGAGCGCACTGCG 769
DB 2570 CGCGCGGAGGCG 2629
QY 770 CGCGCATTCGTGCG 829
DB 2630 GCTGGAAGCG 2689
QY 830 AGCGCGACTGCGCGCTTCCAGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
DB 2690 AGCG 2749
QY 890 AAGTGCCTGCGCACTGCG 949
DB 2750 CGGTGCGCGGTGCG 2809
QY 950 TCGCGCGCGCGCGAGAGCTTCTTGTGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1009
DB 2810 AGCG 2869
QY 1010 CGCTGCGCGCTTGCAGTACG 1069
DB 2870 CG 2929
QY 1070 CGCAGCG 1129
DB 2930 CCTCATGTTGACCG 2986
QY 1130 TCGCGAGCG 1189
DB 2987 CG 3046
QY 1190 GCATCAACG 1249
DB 3047 GAGCG 3106
QY 1250 CTTCT 1255
DB 3107 CGGCT 3112

RESULT 14
US-09-259-821A-1
Sequence 1, Application US/09259821A
Patent No. 6210926
GENERAL INFORMATION:
APPLICANT: LEOPARDI, ROSARIO
TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
FILE REFERENCE: ARCD:317
CURRENT FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 08/690,473
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1

LENGTH: 4257
 TYPE: DNA
 ORGANISM: HERPES VIRUS, TYPE 1
 US-09-821A-1

Query Match 4.6%; Score 77.6; DB 3; Length 4257;
 Best Local Similarity 43.3%; Pred. No. 7.4e-06;
 Matches 548; Conservative 0; Mismatches 694; Indels 24; Gaps 3;

8 ACAAGGTGACCGGACCGCTTCCAGACCGCCAGCCCGGCGCTTCCGACCGCCG 67
 Db AGAGAGAGAGTCCCGCGCTTCCGCGCGCGACCGGCGCGCGCGCGCGCGGCT 1912
 Oy CGTGTGCG 127
 Db AGCG 1972
 Oy 1913 AGCG 1972
 Db 128 CGCAGCG 187
 Oy 1973 CG 2032
 Db 188 ACAAGCTTCTGCAATGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCG 247
 Oy 2033 GCG 2092
 Db 248 GCGTGTGCG 307
 Oy 2093 CGCTGTGCG 2152
 Db 308 CGCAGCG 367
 Oy 2153 CG 2212
 Db 368 TCACG 427
 Oy 2213 CGCAGCG 2269
 Db 428 TGTGTGCG 487
 Oy 2270 TGTGTGCG 2329
 Db 488 CGGTGTCTTTCACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547
 Oy 2330 CG 2389
 Db 548 AGCGCAT-----CACCGACGAGGTGCAAGCTTCTGAGAGCGCGCTTG 589
 Oy 2390 CG 2449
 Db 590 CTGACAGCG 649
 Oy 2450 TGTGTGCG 2509
 Db 650 GCAGCG 709
 Oy 2510 CG 2569
 Db 710 ACTGTGCG 769
 Oy 2570 CG 829
 Db 770 CGCGCATTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
 Oy 2630 GCTGTGCG 2689
 Db 830 AGCG 889
 Oy 2690 AGCG 2749
 Db 890 AAGTGTGCG 949
 Oy 2750 CGGTGTGCG 2809

Oy 950 TGTGTGCG 1009
 Db 2810 AGCG 2869
 Oy 1010 CGTGTGCG 1069
 Db 2870 CG 2929
 Oy 1070 CGCAGCG 1129
 Db 2930 CGCTGTGCG 2986
 Oy 1130 TGTGTGCG 3046
 Db 2987 CG 3046
 Oy 1190 GCATGAGCG 1249
 Db 3047 GCG 3106
 Oy 1250 CCTTCT 1255
 Db 3107 CGGCGT 3112

RESULT 15

US-08-843-659-1
 Sequence 1, Application us/08843659

Patent No. 6218103

GENERAL INFORMATION:

APPLICANT: Leopardi, Roarlio

APPLICANT: Roizman, Bernard

TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS

TITLE OF INVENTION: INHIBITORS OF APOPTOSIS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSES: Arnold, White & Durkee

STREET: P. O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/843,659

FILING DATE: Concurrently Herewith

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: ARSB:519

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4257 base pairs

TYPE: nucleic acid

STANDARDNESS: single

TOPOLOGY: linear

US-08-843-659-1

Query Match

Best Local Similarity 4.6%; Score 77.6; DB 3; Length 4257;

Matches 548; Conservative 0; Mismatches 694; Indels 24; Gaps 3;

Oy 8 ACAAGGTGACCGGACCGCTTCCAGACCGCCAGCCCGGCGCTTCCGACCGCCG 67
 Db 1853 AGAGAGAGAGTCCCGCGCTTCCGCGCGCGACCGGCGCGCGCGCGCGGCT 1912

QY 68 CAGTCTGCGCCCGGCGCCGAGCGCCCGTCTGAGCGCGCGCTTCCACG 127
 Db 1913 ACCGCGCGCGGAGATCTTCCGCGCCCTGAGCGCGCTTCCGCGCGCGCTTCCCG 1972
 QY 128 CGCAGCGCAAGGAGAGATGCGGCTGAGCGGCGCATCGCGCGCGGCGCAAGCGGCGCA 187
 Db 1973 CGGAGGCGAGACCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2032
 QY 188 ACAGCTTCTGCAATGAGCTGACATGTCACGCAAGCGCGCGCATCGGCTGACAGAGAGTGC 247
 Db 2033 GCGCGCAGGCG 2092
 QY 248 GCGTGCCTGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307
 Db 2093 CGCTGCGCGAGGCGCTTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2152
 QY 308 CGCAGCATTCGTGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
 Db 2153 CCGCGAGCG 2212
 QY 368 TCACG 427
 Db 2213 CCGACG 2269
 QY 428 TGCTGCGCTTACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487
 Db 2270 TGGTGTCTATGCG 2329
 QY 488 CGGTGTCTTCTCATGCG 547
 Db 2330 CCG 2389
 QY 548 ACCGCGAT-----CACCGAGAGTGCAGCTTCTGAGAGCGCGCTTGC 589
 Db 2390 CCGCGCTGCGCGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2449
 QY 590 CTGACAGCTGACCG 649
 Db 2450 TGCG 2509
 QY 650 GCAGCGCAGAGCG 709
 Db 2510 CCGCGCGCTTCCG 2569
 QY 710 AGCTGAGCGACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 769
 Db 2570 CGCGCGGAGCG 2629
 QY 770 CCGCATTCGTGCG 829
 Db 2630 GCTCGAGAGCG 2689
 QY 830 AGGCGGACCTGCGCTTCCAGCGMACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
 Db 2690 AGCG 2749
 QY 890 AAGTGGCTTGGCGACCG 949
 Db 2750 CCGTGGCGGTGTGCG 2809
 QY 950 TGCGGCGCGCGCGAGAGCTTTCGTGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1009
 Db 2810 AGCG 2869
 QY 1010 GCGTGGCGCTTGGGATCATGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1069
 Db 2870 CCG 2929
 QY 1070 CCGACCGCAGACGCGCTGCTGACCGCGCAAGAGTACTGAGTGTGAGCGCGCGCGCGCG 1129
 Db 2930 CCTCTATGTGTGACCG 2986

QY 1130 TGCAGAGCGCGCGCGAGCGCGCTGAGCTTCAAGCTGCGCGCGCGCGCGCGCGCG 1189
 Db 2987 CCG 3046
 QY 1190 GCATCAAGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1249
 Db 3047 GCG 3106
 QY 1250 CTTTCT 1255
 Db 3107 CGGCGCT 3112

Search completed: January 20, 2004, 18:00:10
 Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2004, 14:33:57 ; Search time 6062 Seconds

(without alignments)
11303.798 Million cell updates/sec

Title: US-09-808-387-21

Perfect score: 1675
Sequence: 1 atgsgcaacacggtgcacgc.....ctacagcctacagcctacgc 1675

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrc:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	724.8	43.3	189479	9	AC027309	AC027309 Homo sapi
2	724.8	43.3	193960	2	AC022217	AC022217 Homo sapi
3	724.8	43.3	326249	2	AC034210	AC034210 Homo sapi
C 4	499.4	29.8	219695	10	AC122252	AC122252 Mus muscu
C 5	449.2	26.8	238906	2	AC095473	AC095473 Rattus no
6	449.2	26.8	277331	2	AC110151	AC110151 Rattus no
7	409.2	24.4	2207	6	E26374	E26374 Neutalized
8	409.2	24.4	2207	9	U87864	U87864 Homo sapien
9	409.2	24.4	3634	9	AF029729	AF029729 Homo sapi
10	407.6	24.3	4011	9	BC026336	BC026336 Homo sapi
11	399.6	23.9	2445	10	MMNRDRLI	Y15160 Mus musculti
12	396.4	23.7	2210	10	AF400063	AF400063 Mus muscu
13	393.2	23.5	3630	10	AF401228	AF401228 Mus muscu
14	392.2	23.4	1641	6	E26375	E26375 Neutalized
C 15	323.4	19.3	277331	2	AC110151	AC110151 Rattus no
C 16	199.6	11.9	3176	5	AF419159	AF419159 Xenopus l
C 17	186.8	11.2	168481	2	AC025094	AC025094 Homo sapi
C 18	186.8	11.2	203262	9	HSBA416N2	AL121929 Human DNA
19	183.4	10.9	1005	10	MM0271921	AJ271921 Mus muscu
C 20	178.6	10.7	178832	2	AC129799	AC129799 Rattus no
C 21	178.6	10.7	186929	10	AC090657	AC090657 Mus muscu
C 22	178.6	10.7	191825	2	AC132288	AC132288 Mus muscu
C 23	178.6	10.7	226140	2	AC097415	AC097415 Rattus no
C 24	178.6	10.7	233713	2	AC133034	AC133034 Rattus no
C 25	178.6	10.7	282915	2	AC098165	AC098165 Rattus no
26	177.2	10.6	193627	2	AC091427	AC091427 Mus muscu
C 27	170.8	10.2	110000	2	AC132401_1	Continuation (2 of
28	166.4	9.9	1704	9	BC012317	BC012317 Homo sapi
29	166.4	9.9	1733	9	IR2004990	AL889981 Homo sapi
C 30	165.4	9.9	236486	9	AC021188	AC021188 Homo sapi
C 31	158.4	9.5	2207	3	DVNEUR62	U12592 Drosophila
C 32	138.4	8.3	168337	2	AC118956	AC118956 Rattus no
C 33	135.8	8.1	2599	10	AF321278	AF321278 Mus muscu
C 34	134.8	8.0	224371	2	AC132456	AC132456 Mus muscu
35	128.6	7.7	1198	10	MM0271920	AJ271920 Mus muscu
36	127	7.6	2656	3	AY051987	AY051987 Drosophila
37	127	7.6	3396	3	BT003772	BT003772 Drosophila
C 38	127	7.6	123579	2	AC017792	AC017792 Drosophila
C 39	127	7.6	176036	3	AC009258	AC009258 Drosophila
C 40	127	7.6	306651	3	AB003681	AB003681 Drosophila
41	126.6	7.6	3542	3	S62597	S62597 neu-neural1
42	125.4	7.5	3999	3	DROC3HC4ZF	L12218 Drosophila
43	125.4	7.5	3999	3	S62583	S62583 neu-neural1
C 44	119	7.1	46965	2	AC006183	AC006183 Homo sapi
45	115.8	6.9	110000	2	AL831721_0	AL831721 Danilo ter

ALIGNMENTS

RESULT 1
AC027309 189479 bp DNA linear PRI 30-JUL-2002
LOCUS Homo sapiens chromosome 5 clone CTB-7988, complete sequence.
DEFINITION
AC027309
ACCESSION
VERSION
AC027309.3 GI:22004171
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (Bases 1 to 189479)
AUTHORS
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
Direct Submission
JOURNAL
Unpublished

[illegible]

Db 129561 GACGTGCTCCGGGACATCAACGGGCGTCCGGCGGCGCCGCTGCTGTGCGTGCACACACG 129620

Qy 1237 CAGCGCGCTCTGGGCGCTTCTTGGCCGTGGCGGCGCGCGCGTCCGCGGCGCACTGTGCTCTTC 1296

Db 129621 CAGCGCGCTCTGGGCGCTTCTTGGCCGTGGCGGCGCGCGTCCGCGGCGCACTGTGCTCTTC 129680

Qy 1297 GGTACCGCTGCAGTCCAGCGCTGCG 1320

Db 129681 GGTGAGTCCCCCGGCCCCGCGTCCG 123704

RESULT 2

AC022217

LOCUS AC022217 193960 bp DNA linear PRI 09-MAY-2001

DEFINITION Homo sapiens BAC clone RP11-779018 from 5, complete sequence.

ACCESSION AC022217

VERSION AC022217.5 GI:14010943

KEYWORDS HTG.

SOURCE

ORGANISM Homo sapiens (human)

Bufo vulgaris, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Mammalia, Eutheria, Primates, Catarrhini, Homiidae, Homo.

REFERENCE

AUTHORS 1 (bases 1 to 193960)
TITLE Sulston, J.B. and Waterston, R.
JOURNAL Toward a complete human genome sequence
MEDLINE Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 95063792

REFERENCE

AUTHORS 2 (bases 1 to 193960)
TITLE Kruczkowski, S., Stoneking, T., Hawking, M., Strommat, C. and Le, T.
JOURNAL The sequence of Homo sapiens BAC clone RP11-779018
REFERENCE Unpublished
AUTHORS 3 (bases 1 to 193960)
TITLE Waterston, R.H.
JOURNAL Direct Submision
AUTHORS Submitted (09-MAY-2001) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 193960)
TITLE Waterston, R.
JOURNAL Direct Submision
AUTHORS Submitted (09-MAY-2001) Department of Genetics, Washington
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 9, 2001 this sequence version replaced gi:11245628.

COMMENT

Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu
Summary Statistics
Center project name: H_NH0779018

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:


```

Db      3390 GAGAGCGCGCTTGTGACGACGCTGACGCGCGCGCTGACGACGCGCGCTTCAAGCGCC 3449
Qy      637 TGCCTGCGCGCCAGACAGCCAGACCGCGGCACTTGTGACACAGAGCTTCAAGAACAC 696
Db      3450 TGCCTGCGCGCCAGACAGCCAGACCGCGGCACTTGTGACACAGAGCTTCAAGAACAC 3509
Qy      697 CAGGTGTGTGACCAAGCTGTGACCGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
Db      3510 CAGGTGTGTGACCAAGCTGTGACCGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3569
Qy      757 GACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 816
Db      3570 GACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3629
Qy      817 CCGGCGCTACTGTGAAGCGCGACCTGTGCGCTTCCAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 876
Db      3630 CCGGCGCTACTGTGAAGCGCGACCTGTGCGCTTCCAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 3689
Qy      877 TCGGCGCGACCGCAAAAGTGTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 936
Db      3690 TCGGCGCGACCGCAAAAGTGTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3749
Qy      937 TCGGAGCGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 996
Db      3750 TCGGAGCGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3809
Qy      997 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1056
Db      3810 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3869
Qy      1057 AACGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1116
Db      3870 AACGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3929
Qy      1117 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176
Db      3930 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3989
Qy      1177 GACGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1236
Db      3990 GACGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4049
Qy      1237 CAGGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1296
Db      4050 CAGGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4109
Qy      1297 GGTACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db      4110 GGTGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4133

RESULT 3
AC034210/c 36249 bp DNA linear HTG 22-JUN-2000
LOCUS      AC034210.2 GI:8655959
DEFINITION Homo sapiens chromosome 5 clone CTC-537115, *** SEQUENCING IN
PROGRESS ***, 100 unordered pieces.
ACCESSION AC034210.2
VERSION    AC034210.2
KEYWORDS   HTG; HTGS; PHASE1.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 36249)
AUTHORS   DOE Joint Genome Institute.
TITLE     Sequencing of Human Chromosome 5
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 36249)
AUTHORS   DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

```

COMMENT

```

On Jun 22, 2000 this sequence version replaced gi:7417687.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 391595
Center clone name: CIT-HSPC_537115
-----
Summary Statistics
Consensus quality: 24405 bases at least Q40
Consensus quality: 283546 bases at least Q30
Consensus quality: 293437 bases at least Q20
Estimated insert size: 153000; pulse field gel estimation
Estimated insert size: 316349; sum-of-contigs estimation
Quality coverage: 6.25 in Q20 bases; pulse field gel estimation
Quality coverage: 3.02 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 100 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1159: contig of 1159 bp in length
* 1160 1259: gap of unknown length
* 1260 2490: contig of 1231 bp in length
* 2491 2590: gap of unknown length
* 2591 3801: contig of 1211 bp in length
* 3802 3901: gap of unknown length
* 3902 5076: contig of 1175 bp in length
* 5077 5176: gap of unknown length
* 5177 6194: contig of 1018 bp in length
* 6195 6294: gap of unknown length
* 6295 7358: contig of 1064 bp in length
* 7359 7459: gap of unknown length
* 7459 9384: contig of 1926 bp in length
* 9385 9484: gap of unknown length
* 9485 10606: contig of 1122 bp in length
* 10607 10706: gap of unknown length
* 10707 11830: contig of 1124 bp in length
* 11831 11930: gap of unknown length
* 11931 13117: contig of 1187 bp in length
* 13118 13217: gap of unknown length
* 13218 14757: contig of 1540 bp in length
* 14758 14857: gap of unknown length
* 14858 15960: contig of 1103 bp in length
* 15961 16060: gap of unknown length
* 16061 17225: contig of 1165 bp in length
* 17226 17325: gap of unknown length
* 17326 18483: contig of 1158 bp in length
* 18484 18583: gap of unknown length
* 18584 19725: contig of 1142 bp in length
* 19726 19825: gap of unknown length
* 19826 20844: contig of 1019 bp in length
* 20845 20945: gap of unknown length
* 20946 22436: contig of 1492 bp in length
* 22437 22536: gap of unknown length
* 22537 23637: contig of 1001 bp in length
* 23638 25102: gap of unknown length
* 25103 25202: contig of 1465 bp in length
* 25203 26114: gap of unknown length
* 26115 26514: gap of unknown length
* 26515 27649: contig of 1135 bp in length
* 27650 27749: gap of unknown length
* 27750 28974: contig of 1225 bp in length
* 28975 29074: gap of unknown length
* 29075 30458: contig of 1384 bp in length
* 30459 31769: gap of unknown length
* 31769: contig of 1211 bp in length

```

* 31770 31859: gap of unknown length
* 31870 33499: contig of 1630 bp in length
* 33500 33599: gap of unknown length
* 33600 34981: contig of 1382 bp in length
* 34982 35081: gap of unknown length
* 35082 36384: contig of 1303 bp in length
* 36385 36484: gap of unknown length
* 36485 37903: contig of 1419 bp in length
* 37904 38003: gap of unknown length
* 38004 39584: contig of 1581 bp in length
* 39585 39684: gap of unknown length
* 39685 41198: contig of 1514 bp in length
* 41199 41299: gap of unknown length
* 41299 42600: contig of 1301 bp in length
* 42600 42700: gap of unknown length
* 42700 44201: contig of 1502 bp in length
* 44202 44301: gap of unknown length
* 44302 45882: contig of 1581 bp in length
* 45883 45982: gap of unknown length
* 45983 47265: contig of 1283 bp in length
* 47266 47365: gap of unknown length
* 47366 48453: contig of 1088 bp in length
* 48454 48553: gap of unknown length
* 48554 49848: contig of 1295 bp in length
* 49849 49949: gap of unknown length
* 49949 51156: contig of 1208 bp in length
* 51157 51256: gap of unknown length
* 51257 52274: contig of 1018 bp in length
* 52275 52374: gap of unknown length
* 52375 53924: contig of 1550 bp in length
* 53925 54024: gap of unknown length
* 54025 55407: contig of 1383 bp in length
* 55408 55507: gap of unknown length
* 55508 56716: contig of 1109 bp in length
* 56717 58308: gap of unknown length
* 58309 58408: contig of 1592 bp in length
* 58409 59912: gap of unknown length
* 59912 60012: contig of 1504 bp in length
* 60013 61639: contig of 1627 bp in length
* 61640 61739: gap of unknown length
* 61740 63207: contig of 1468 bp in length
* 63208 63307: gap of unknown length
* 63308 65127: contig of 1820 bp in length
* 65128 65227: gap of unknown length
* 65228 67091: contig of 1864 bp in length
* 67092 67191: gap of unknown length
* 67192 68762: contig of 1571 bp in length
* 68763 68862: gap of unknown length
* 68863 70597: contig of 1735 bp in length
* 70598 72482: contig of 1785 bp in length
* 72483 72582: gap of unknown length
* 72583 74499: contig of 1917 bp in length
* 74500 74599: gap of unknown length
* 74600 76545: contig of 1946 bp in length
* 76546 76645: gap of unknown length
* 76646 78780: contig of 2135 bp in length
* 78781 78880: gap of unknown length
* 78881 80540: contig of 1660 bp in length
* 80541 80640: gap of unknown length
* 80641 82351: contig of 1711 bp in length
* 82352 82451: gap of unknown length
* 82452 84871: contig of 2420 bp in length
* 84872 84972: gap of unknown length
* 84973 87283: contig of 2312 bp in length
* 87284 87383: gap of unknown length
* 87384 89384: contig of 2001 bp in length
* 89385 89484: gap of unknown length
* 89485 91357: contig of 1873 bp in length
* 91358 91457: gap of unknown length
* 91458 93682: contig of 2225 bp in length
* 93683 93782: gap of unknown length

* 93783 95479: contig of 1697 bp in length
* 95480 95579: gap of unknown length
* 95580 97605: contig of 2026 bp in length
* 97606 97705: gap of unknown length
* 97706 99306: contig of 1601 bp in length
* 99307 99406: gap of unknown length
* 99407 101015: contig of 1609 bp in length
* 101016 101115: gap of unknown length
* 101116 102573: contig of 1458 bp in length
* 102574 102673: gap of unknown length
* 102674 105086: contig of 2413 bp in length
* 105087 105186: gap of unknown length
* 105187 107178: contig of 1992 bp in length
* 107179 107278: gap of unknown length
* 107279 110942: contig of 3664 bp in length
* 110943 112727: contig of 1685 bp in length
* 112728 112827: gap of unknown length
* 112828 115272: contig of 2445 bp in length
* 115273 115372: gap of unknown length
* 115373 117878: contig of 2506 bp in length
* 117879 117978: gap of unknown length
* 117979 121421: contig of 3443 bp in length
* 121422 121521: gap of unknown length
* 121522 123917: contig of 2396 bp in length
* 123918 124017: gap of unknown length
* 124018 126854: contig of 2837 bp in length
* 126855 126954: gap of unknown length
* 126955 130347: contig of 3393 bp in length
* 130348 130447: gap of unknown length
* 130448 132718: contig of 2271 bp in length
* 132719 132818: gap of unknown length
* 132819 136122: contig of 3304 bp in length
* 136123 136222: gap of unknown length
* 136223 138733: contig of 2511 bp in length

Query Match 43.3%; Score 724.8; DB 2; Length 326249;
Beet Local Similarity 98.4%; Pred. No. 7.7e-70;
Matches 732; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 577 GAGAGCCCTTGGCTGACGACCTGACCGCCGCGGCGCTGACGACCGCCGCTTACGCGCC 636
DB 268092 GAGAGCCCTTGGCTGACGACCTGACCGCCGCGGCGCTGACGACCGCCGCTTACGCGCC 268033
QY 637 TGCTGCGCGCCAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 696
DB 268032 TGCTGCGCGCCAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 267973
QY 697 CAGGTGTGCGCAGAGCTGGGCGACCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
DB 267972 CAGGTGTGCGCAGAGCTGGGCGACCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 267913
QY 757 GAGCG 816
DB 267912 GAGCG 267853
QY 817 CCGCGCTACTGAGAGCGCGAGCTGGCGCTTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 876
DB 267852 CCGCGCTACTGAGAGCGCGAGCTGGCGCTTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 267793
QY 877 TCGGCGCAGCGCAGAAAGTGGCTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 936
DB 267792 TCGGCGCAGCGCAGAAAGTGGCTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 267733
QY 937 TCGGAGCG 996
DB 267732 TCGGAGCG 267673
QY 997 GCGGCG 1056
DB 267672 GCGGCG 267613
QY 1057 AACGAGCTGCCCG 1116


```

Db      267612 AACGAGCTGCCCCGACCCAGACGCGCTCTGACGCGCAAGAGTACTGGGTGGCG 267553
QY      1117 CCGGCGGGGCGCCCTGCGGAGGCGGCGGAGCGGCTCAAGTTCAACGCTGCGGCGCGCGCG 1176
Db      267552 CCGGCGGGGCGCCCTGCGGAGGCGGCGGAGCGGCTCAAGTTCAACGCTGCGGCGCGCGCG 267493
QY      1177 GACGTCTCTGGGCAATCAACGAGGCGGCGGCGGCGGCGGCTGCTGAGCTGACACAGC 1236
Db      267492 GACGTCTCTGGGCAATCAACGAGGCGGCGGCGGCGGCGGCTGCTGAGCTGACACAGC 267433
QY      1237 CAGGCGCTCTGGGCGCTTCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCTTC 1296
Db      267432 CAGGCGCTCTGGGCGCTTCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCTTC 267373
QY      1297 GGTACCTGTGAGTCCAGCCTTGGC 1320
Db      267372 GGTGAGTCCCGGCGCGCGTGGC 267349

```

RESULT 4

```

AC122252/c      219695 bp      DNA      linear      ROD 29-MAY-2003
LOCUS      Mus musculus chromosome 17 clone RP23-170E7, complete sequence.
DEFINITION      AC122252
ACCESSION      AC122252
VERSION      AC122252.4 GI:31126889
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE      1 (bases 1 to 219695)
AUTHORS      Wilson, R.K.
TITLE      The sequence of Mus musculus clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 219695)
AUTHORS      McPherson, J.D. and Waterston, R.H.
TITLE      McPherson, J.D. and Waterston, R.H.
JOURNAL      Direct Submission
              Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE      3 (bases 1 to 219695)
AUTHORS      McPherson, J.D. and Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (18-MAR-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE      4 (bases 1 to 219695)
AUTHORS      Wilson, R.K.
TITLE      Direct Submission
JOURNAL      Submitted (29-MAY-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
COMMENT      On May 29, 2003 this sequence version replaced gi:29029447.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu

```

```

----- Project Information -----
Center project name: M_BA0170E07

```

FEATURES

```

source      1. 219695
              Location/Qualifiers
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /DB_xref="taxon:10090"
              /chromosome="17"
              /clone="RP23-170E7"

```

```

BASE COUNT      54661 a 54739 c 54797 g 55498 t
ORIGIN

```

```

Query Match      29.8%; Score 499.4; DB 10; Length 219695;
Best Local Similarity 82.3%; Pred. No. 1,3e-45;
Matches 597; Conservative 0; Mismatches 101; Indels 27; Gaps 1;

```

```

QY      575 TGAAGAGCGCTTCTGCTGACAGCTGACGCGCGCGCTTCAAGCGAGCGCGCTTCAAGCG 634
Db      204239 TAGAAGAGCGCTTCTGCTGACAGCTGACGCGCGCGCTTCAAGCGAGCGCGCTTCAAGCG 204180
QY      635 CTTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 694
Db      204179 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204120
QY      695 ACCAGGTGAGCGCAAGCTGAGCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 754
Db      204119 ACCAGGTGAGCGCAAGCTGAGCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204076
QY      755 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
Db      204075 -----GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204027
QY      815 CCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 874
Db      204026 CCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203967
QY      875 TGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 934
Db      203966 TGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203907
QY      935 TCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 994
Db      203906 TCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203847
QY      995 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1054
Db      203846 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203787
QY      1055 CCAAGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1114
Db      203786 CATCCAGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203727
QY      1115 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1174
Db      203726 CTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203667
QY      1175 GCGAGCTGCTCTGAGGCAATCAACGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1234
Db      203666 GCGAGCTGCTCTGAGGCAATCAACGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203607
QY      1235 CCGAGGCGCGCTGAGGCGCTTCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1294
Db      203606 CACAGGCGCGCTGAGGCGCTTCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 203547
QY      1295 TCGGT 1239
Db      203546 TCGGT 203542

```

RESULT 5

```

AC095473/c      238906 bp      DNA      linear      HTG 09-MAY-2003
LOCUS      Rattus norvegicus clone CH230-713, WORKING DRAFT SEQUENCE.
DEFINITION      AC095473
ACCESSION      AC095473
VERSION      AC095473.6 GI:30467680
KEYWORDS      HTG, HTGS_PHASE2, HTGS_DRAFT, HTGS_FULLTOP.
SOURCE      Rattus norvegicus
              Rattus norvegicus (Norway rat)
              Rukayocsa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```

```

REFERENCE      1 (bases 1 to 238906)
AUTHORS      Muzny, D., Matle, M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
              Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
              Anyalebechi, V., Ayegbi, A., Ayodeji, M., Baca, B., Baden, H.,
              Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F.,
              Biewick, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
              Bryant, N., Bunney, C., Burch, F., Burrell, K., Calderon, B.,

```

Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Dvali, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guvarra, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hui, S., Hume, J., Idubrid, D., Jackson, A., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwie, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzulewa, L., Louisedge, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milsosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervels, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokweli, O., Okunolu, G., Olarunpasegun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfennich, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shaeman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sleson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Swack, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, D., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooder, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 238906)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238906)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24941075.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCRH
Center clone name: CH230-713
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 227630 bases at least Q40
Consensus quality: 229694 bases at least Q30
Consensus quality: 231141 bases at least Q20
Estimated insert size: 238998; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 238906: contig of 238906 bp in length.
Location/Qualifiers
1. 238906
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-713"
1. 1595
/note="wgs contig"

BASE COUNT 57521 a 59144 c 57758 g 57899 t 6584 others
ORIGIN
Query Match 26.8%; Score 449.2; DB 2; Length 238906;
Best Local Similarity 87.8%; Pred. No. 3e-40;
Matches 490; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

24 CCTGCCAGACCGAGCCCAACCGGCGCGCTCTGAGCAACCGGCGGCTGCGGCGCCCGG 83
Db 81151 CCCACAGATTCAGACCTCTCTGAGCGCTCTGAGCAACCGGCGGCTGAGCGCCCGG 81092

84 CCCGAGCGACGCGCGCTCTGAGCGCGCGCGCTTCCAGCGCGGCAAGGCAA 143
Db 81091 CCCGAGCGCGAGCGGCTCTGAGCGCGCGCGCTTCCAGCGCGGCAAGGCAA 81032

144 GAACTGCGGCTGAGCGGCACTGCGCGCGCGGCAACGCGGCAAGCTTTCGCAATG 203
Db 81031 GAACTGCGGCTGAGCGGCACTGCGCGCGCGGCAACGCGGCAAGCTTTCGCAATG 80972

204 CGTCAAGCTTACCGACGCGCGCGCTGAGCGGCGGCTGAGCGGCGGCTGAGGCG 263
Db 80971 AGTCAAGCTTACCGACGCGCGCGCTGAGCGGCGGCTGAGCGGCGGCTGAGGCG 80912

264 CGTGGCGCTGAGCGGCGGCGGCTGAGCGGCGGCTGAGCGGCGGCTGAGCGGCG 323
Db 80911 TGTGCGCTGAGCGGCGGCGGCTGAGCGGCGGCTGAGCGGCGGCTGAGCGGCG 80852

324 CATGAGCGCCAGAGCATCCCAAGTACGCTGCGCGGCACTGAGCGGCGGCGGCTA 383
Db 80851 CATGAGCGCCAGAGCATCCCAAGTACGCTGCGCGGCACTGAGCGGCGGCGGCTA 80792

384 CTGGGCGCAAGGCACTGCGCGGCAAGTACGCTGCGCGGCACTGAGCGGCGGCGGCT 443
Db 80791 CTGGGCGCAAGGCACTGCGCGGCAAGTACGCTGCGCGGCACTGAGCGGCGGCGGCT 80732

444 CGACCGCGACGCGCGCTGTTCTTCAAGCGTGAACGAGCGGCGGCTTTCGCACTG 503
Db 80731 TGAATCGTCAAGCGGCGGCTTTCGCAAGCGTGAACGAGCGGCGGCTTTCGCACTG 80672

504 CGGCGTGGCGGCGGCGGCGGCGGCTGAGCGGCGGCTGAGCGGCGGCTGAGCGGCG 563

Db	80671	CGGCTGGCCCTGGAGAGGCCCACTCTGGGCACACTATGCAACGTTCATTGTCATCAGCACGA	806151
Oy	564	GGTGCAGCTTCTGAGAG	581
Db	80611	GGTGCAGCTGCTGGGTAG	80594
RESULT 6			
AC110151		277311 bp	DNA linear HTG 22-SEP-2002
DEFINITION	Rattus norvegicus clone CH230-47B19,	*** SEQUENCING IN PROGRESS	
LOCUS	*** 3 unordered pieces.		
ACCESSION	AC110151		
VERSION	AC110151.4	GI:23195250	
SOURCE	HTG, HTGS, PHASE1, HTGS DRAFT, HTGS_ENRICHED.		
ORGANISM	Rattus norvegicus (Norway rat)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 277311)		
AUTHORS	Mueny D.Marie, Metzker M.Lee, Abramzon S., Adams C., Alder J., Allen C., Allen H., Albrooke S., Amin A., Angiano D., Anulescchi V., Aoyagi I., Ayodeji M., Baca E., Baden H., Baldwin D., Bandaranaike D., Barber M., Barnstead M., Behnsted F., Blawie O.K., Blair J., Blankenburg K., Blyth P., Brown M., Bryant N., Buhay C., Burch P., Butrell K., Calderon E., Cardenas V., Carter K., Cavazos I., Cesar H., Center A., Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J., Cleaveland C., Cockrell R., Cox C., Coyle M., Cree A., pSouza L., Davila M.L., Davis C., Davy-Carroll U., De Anda C., Dederich D., Delgado O., Denison S., Deramo C., Ding Y., Din H., Divya K., Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Evans K., Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P., Fraser C.M., Gabisi A., Garcia R., Garcia A., Garner T., Garza M., Georgegeorgis B., Geer K., Gill R., Grady M., Guerra W., Guevara M., Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K., Harvey Y., Hawlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M., Hollins B., Howells S., Huliyk S., Hume J., Ididatid D., Jackson A., Jackson L., Jacod L., Jiang H., Johnson B., Johnson R., Jolyet A., Kapatchy S., Kelly S., Kelly S., Khan Z., King L., Kovac J., Kowls C., Kreft C.L., Lebow H., Levam L., Lewis L., Li Z., Liu J., Liu J., Liu Y., Liu Y., London P., Longacre S., Lopez J., Lorensheva L., Louised H., Lozada R.J., Lu X., Ma J., Maheshwari M., Mahindratne M., Mamoud M., Malloy K., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Martinez E., Mahoney S., McLeod M.P., McNella T.Z., Meenen B., Milosavljevic A., Miner G., Minja E., Montemayor J., Morgan M., Morris K., Morris S., Mundasa M., Murphy M., Natr L., Nankerris C., Neal D., Newton N., Nguyen N., Norris S., Nwokelamen O., Okwonu G., Olairumagoun A., Pal S., Parke K., Paeteren S., Paul H., Perez A., Perez L., Pfannkoch C., Plapper F., Poindexter A., Popovic D., Primus E., Pu L.C., Puzo M., Quiroz J., Rachlin E., Reeves K., Reiter M.A., Reigh R., Kelly B., Kelly M., Ren Y., Reuter M., Richards S., Riggs F., Rivers C., Rodney T., Rojao A., Rose M., Rose R., Ruiz S.J., Sanders W., Savery G., Scherer S., Scott G., Shatman S., Shen H., Shetty J., Shvartsbeyn A., Sisson I., Slitter C.D., Smjs D., Sneed A., Sodergren E., Song X.-Z., Sokela R., Soea J., Steidle M., Strong R., Sutton A., Svetek A., Taboi P., Taylor C., Taylor T., Thomas N., Thomas S., Tinney A., Trejos Z., Umami K., Valas R., Vera V., Williams D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren R., Wei X., White F., Williams G., Willson R., Wlarczyk R., Wooden H., Wolley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu P., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.		
TITLE	Direct Submission		
JOURNAL	Unpublished		

```

REFERENCE
AUTHORS      2 (bases 1 to 277331)
TITLE        Worley K.C.
JOURNAL      Direct Submission
              Submitted (10-FEB-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              3 (bases 1 to 277331)
REFERENCE    Rat Genome Sequencing Consortium.
AUTHORS      Direct Submission
              Submitted (22-SEP-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              On Sep 19, 2002 this sequence version replaced gi:21743237.
COMMENT      The sequence in this assembly is a combination of BAC based reads
              and whole genome shotgun sequencing reads assembled using Atlas
              (http://www.hgsc.bcm.tmc.edu/projects/atlantis/). As a result, the
              sequence may extend beyond the ends of the clone and there may be
              contigs that consist entirely of whole genome shotgun sequence
              reads. Both end sequences and whole genome shotgun sequence only
              contigs will be indicated in the feature table.

              ----- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu

              ----- Project Information
              Center project name: GRPU
              Center clone name: CH230-47B19

              ----- Summary Statistics
              Assembly program: Phrap; version 0.990329
              Consensus quality: 231941 bases at least Q40
              Consensus quality: 236540 bases at least Q40
              Consensus quality: 239843 bases at least Q20
              Estimated insert size: 252686; sum-of-contigs estimation
              Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

              -----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drafc\_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      13319: contig of 13319 bp in length
*
* 13220  13319: gap of unknown length
*
* 13320  266558: contig of 253239 bp in length
*
* 26659  266658: gap of unknown length
*
* 26659  277331: contig of 10673 bp in length.
*
* location/Qualifiers
*
* 1. 277331
*   /organism="Rattus norvegicus"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:10116"
*   /clone="CH230-47B19"
*
* 416. 1276
*   /note="clone boundary"
*   clone_end:5p6
*   site:Scori
*
* end sequence: BH356625"
*
* misc_feature
*   36181.. 37518
*   /note="wgs contig"
*
* misc_feature
*   41065..42752
*   /note="wgs contig"
*
* misc_feature
*   83582.. 86052
*   /note="wgs contig"
*
* misc_feature
*   106318.. 107848
*   /note="wgs contig"
*
* misc_feature
*   137921.. 139155
*   /note="wgs contig"

```

```

misc_feature      complement (245796..246299)
                  /note="clone_boundary
                  clone_end:T7
                  site:ECORI
                  end_sequence:BH356621"
misc_feature      265182..265558
                  /note="wgs_end_extension
                  clone_end:T7"

BASE COUNT      60560 a 61304 c 60133 g 60025 t 35309 others
ORIGIN
Query Match      26.8%; Score 449.2; DB 2; Length 277331;
Best Local Similarity 87.8%; Pred. No. 2.9e-40;
Matches 490; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY      24 CTTGCCAGACCCGAGCCACCGGCGCGCTCTTGCCACCCCGGCGCTGCGCGCCCG 83
Db      38624 CCCACAGACTCCAGCCCTCTGCAAGCTCTTGCCACCCCGGCGCTGCGCGCCCG 38683
QY      84 CCCGAGGAGCGCCCGGCTCTGCGGAGGCGCGCGCTTCCACCGCGAGGCGCAAGGCGA 143
Db      38684 CCCGAGGAGCGAGCGGCTCTGCGGAGGCGCGCGCTTCCACCGCGAGGCGCAAGGCGCA 38743
QY      144 GAAAGTGGGCTGGACGGCCACTGCGCGCGGCGCAAGGCGCAAGCTTTGCAATGG 203
Db      38744 GAAAGTGGGCTGTAAAGGCTGCTGCGAGGCGCGCAAGGCGCAAGCTTTGCAAGCG 38803
QY      204 CGTCACTTCAACGAGCGGCGCATCCGCTGTAGAGAGAGTGGCGCTGCGCTGTAGG 263
Db      38804 AGTACCTTCAACGAGGCGCGCATTCGCTGTAGAGAGAGTGGCGCTGCGCTGTAGG 38863
QY      264 CGTGGCGCTGCTGAGAGCGCGCGCTGCGCTTCTGCTTCAACCGCGCAAGTCCGTGCT 323
Db      38864 TGTCGCTGCTGCTGAGAGCGCGCGCTGCGCTTCTGCTTCAACCGCGCAAGTCCGTGCT 38923
QY      324 CATAGAGGCGCCAGAGCAATCCCAAGTAGAGCTGCGCGGAGCTGTGTCAGCGGCGCGGCTA 383
Db      38924 CATAGAGCAACAGAGATATCCCAAGTAGAGCTGCGCGGAGCTGTGTCAGCAAGCTGTGATA 38983
QY      384 CTGGGCGCAAGGCACTGCGCGGAGAACCTGCGCTGCGCGCAAGCGTGTGCTTACTGAGG 443
Db      38984 CTGGGCGCAAGGCGCTGCGCGGAGAACCTGCGCGCTGCGCGCAAGCGTGTGCTTACTGAGG 39043
QY      444 CGACCGCGAGCGCGCGCTGTTTCAAGCGTGAAGAGGCGCGGCTGCTTCCACTG 503
Db      39044 TGATCGTCAAGCGTGCCTTCTTATAGTGTCAAGTGAAGTGGCGAACAAGTGTGCTTCCACTG 39103
QY      504 CGGCGTGGCGGTGGGCGGCGCGCTGAGGCGCTGATGATGTATACGCAATCAACGAGCA 563
Db      39104 CGGCGTGGCGGTGGGAGGCGCGCACTGTGGGCACTGATGATGTATAGCAATCAACGAGCA 39163
QY      564 GGTGCAAGCTTCTGAGAG 581
Db      39164 GGTGCAAGCTTGTGGTAG 39181

RESULT 7
LOCUS      E26374
DEFINITION Neutralized protein, polynucleotide encoding said protein and
ACCESSION E26374
VERSION    E26374.1 GI:13025070
KEYWORDS   JP 199137257-A/1.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 2207)
AUTHORS    Motomi,N., Hideo,N., Mitsuhiro,Y. and Hideyuki,S.
TITLE      Neutralized protein, polynucleotide encoding said protein and
           antibody recognizing said protein
           Patent: JP 199137257-A 1 25-MAY-1999;
           SUMITOMO ELECTRIC IND LTD

```

```

COMMENT          OS      Unidentified
                  RN      JP 199137257-A/1
                  PD      25-MAY-1999
                  PR      14-NOV-1997 JP 1997313211
                  PX      MOTOMI NAKADA,HIDEO NAKAMURA,MITSUHIRO YOSHIDA,HIDEYUKI SAVA
                  PI      C12N15/09,C07K14/47,C07K16/18,C12P21/02,C12Q1/68,G01N33/53//
                  PC      (C12N15/09,C12R1:91),(C12P21/02,C12R1:19),C12N15/00,(C12N15/00,PC
                  C12R1:91)
                  CC      Strandedness: Double;
                  CC      Topology: Linear;
                  FH      Key
                  FT      source
                  FT      Location/Qualifiers
                  source      1..2207
                             /organism="unidentified"
                             /mol_type="genomic DNA"
                             /db_xref="taxon:32644"
                             /db_872 c 673 g 322 t

BASE COUNT      340 a 872 c 673 g 322 t
ORIGIN
Query Match      24.4%; Score 409.2; DB 6; Length 2207;
Best Local Similarity 56.6%; Pred. No. 2.7e-35;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY      99 GGTCTGGGAGAGGCGCGCGCTTCCACCGCGAGGCCAAGAGCAAGCTGGCTGGA 158
Db      581 GCTCCAGGACCGCGCTGCTTCCACCGCGAGCCAGAGGCTCCAGATCTCATGGA 640
QY      159 CGGCACTGCGCGCGGCGCACAGCGGCAACACTTGTGAATGGGTCAAGTTACGCA 218
Db      641 CCTCAGCCACAGAGCTGTCAAGAGGAGGCGAGCACTTGTGAAGGCCATCCTTACAGCA 700
QY      219 GCGGCGCATCGGCTGTACAGAGAGGTGCGGCTGCGCTGTGGAGCGGCGCTGAGCTG 278
Db      701 CGCGCGGCTCTCATCTACAGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 760
QY      279 GAGCGGCGCTGCGCTTCCGCTTCAACCGCGAGATCGTGTCTATAGCGCCACGA 338
Db      761 GAGCGGCGCTGCGCTTCCGCTTCAACCGCGAGATCGTGTCTATAGCGCGCTTCAACCG 820
QY      339 CATCCCAAGTAAAGCTGCGCGGAGCACTGTGTCAAGCGCGGCTTACTGGGCGCAAGGACT 398
Db      821 GCTGCCAAGTAAAGCTGCGCGGAGCACTGTGTGTCCAGAGTGGCTTGTGGGCGCAAGGCGCT 880
QY      399 GCCGAGAACTGAGCGCTGCGGAGCAAGGAGTGTGCTTACTGTGAGCGCGCGCG 458
Db      881 GCTGAGAGATTTGCAATGAGGCAATCATCTGCACTTGTGGGTGAGCAAGAGGCGCG 940
QY      459 CGTGTCTACAGCGTGAACAGCGGAGCGGAGTCTTTCACCTGCGGCGTGGCGG 518
Db      941 TGCTTTCACCGCATCAAGCACTCGGCTGTATAGTGTTCCTTCAAGCGGAGTCCGACGGC 1000
QY      519 CGGCGCGCTTGGGCGCTCATTTGATGTCTACGGGATCAACGAGAGTGAAGCTTCTGGA 578
Db      1001 GAAACCGCTTGGGCGCTGTGTGAGAGCTTCAAGGCTCAACGCGGCGCTTCAAGGCTTGA 1060
QY      579 GAGGCGCTTGTGAGCAAGCTGAGCGCGCGGCGCTTCAAGCGAGGCGGCTTCAAGGCTG 638
Db      1061 TAGGAGCTG-----TGTCTCCGAGACTGTTCGCGGCGGCGCTTCAAGGCTTCA 1111
QY      639 CTTGGCGCCAGAGCAAGCGCGGCAACTTTCAGCAAGAGTGTGAGAAACAACA 698
Db      1112 GCGGCGGCGCTGCGCTGCGGCGGAGGAGCAAGCGCGGCTTCTCGGTG----- 1160
QY      699 GGTGTGAGCAAGTGGGCGCACTTGGCGCTGGGCGGCGCGCGGCGGCGGCGGCGG 758
Db      1161 -----AGCTTATGCACTTCAAGCTGTGCGGAGGCGGAGCGGCGGCGGCGGCGG 1210
QY      759 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 818

```

Db	1211	GGCGCGCGGCTGCGCCCATCCCGCAGAAACTCACTGAATCGGAGGACAAGCGCGCGCTGCC	1270
Qy	819	GGCGCTACTGAGAGGCCCAACTCTGGCTTTCACAGCAACAAGCGAGGCCG---ACGTGAAGCT	875
Db	1271	GGCGCACACTCGAACGGCGCACTCTGGCTTTCACAGCCCTCGCGCGCGCGCGCAAGTCCGCAT	1330
Qy	876	GTCGCGCGAATCGCAAAAGTGGCTTCGCAACCGCGGCCGGAAGCGCGCGCAAGCTGGTCTT	935
Db	1331	CTTCGAGAGAGAGACGGTGGCGGCCCTGGAGAACACAGGCGCCACAGAGCGCGCTGTCTT	1390
Qy	936	CTCCGAGCGCCCGCTGCGCGGCCCGCGCAGAGAGCTCTTCGTGGAGGTGGGCGCTCCGAGGCT	995
Db	1391	CACCAAGCGGCCCGTGGCGCTGGCGCGAGAACCATCTTCGTCAAAGTCAAGCGCTCGGGTGG	1450
Qy	996	GGCGCGCGCCCGCGCGCGCTTGAGCTTTCGAGCATCAAGTCTGGACGACCCGCGCTGACAGCC	1055
Db	1451	CGCGCGCGCGCGCGCGCTGTGCTTCGAGGTCAACCAAGTGCAGACCCCGGCAAGCTGGCGCC	1510
Qy	1056	CACGAGCTGGCGCGCGCCGACCCAGAACGCGCTGCTGAACCGCAAAAGTACTGGGTGGTGGC	1115
Db	1511	GGCGGACCTGCTCTTTCAGCGCCCTTGAGGCCCTGTGTGAGACGCAAGAAATCTGGGCGGTGTG	1570
Qy	1116	GGGCGCGCGCGCGCGCGCGAGCGCGCGCGCGAGCGCGCTGACTTTCAGCTGCGCGCGCGCGCG	1175
Db	1571	CGCGGTGCGCGCGCGCGCGCGCTTCACACAGCGCGGACATCTGGGCGCTGGTGTGACACGCGAAG	1630
Qy	1176	CGACGTGCTCTTGAGCATCAACAGGAGCGTCCGCGCGCGCGCGCGCTGTGTGGTGTGACACAC	1235
Db	1631	CGAGCTGACCTCAAGCAACAAATGGCGCGCGCGCGCGCATGACAGCTGTGCGTGGAGCGCTTC	1690
Qy	1236	GCAAGCGCTCTGGGCGCTTCTTGCGCGCGTGGCGCGCGCGCGCGCTGGCGAGCTGGCTCTCT	1295
Db	1691	GCAGCGCGCTTGGATGTCTCTTCGCGCTTCACAGGACCAATCAAGCAATCCGCATCTTCGG	1750
Qy	1296	CGGTAACTCTGAGTCCAGCGCCGTGGGACACAGACTCAATCAGGGTCCCTCAG-----	1346
Db	1751	CTCCACTAATCTGGCGCGAGCGGGGTATCCCTCATCTCCCTGCTCTCCCTGCGCTCAAGCC	1810
Qy	1347	-----CGGCTCCACAGACGATAGTGAATTACAGATATGACCTTCAGTGTCAACCACTC	1397
Db	1811	AACCTCGGCCAAGTGGCTCTGGGCAAGCGCGCTGTCTGAACCCCTTGCTCAACAGTGAAGCTC	1870
Qy	1398	CTCTCGGCATCTGAGTATCTCTGGTGAAGGCGGCCCAAGCTTCCCGCTGAAGCCCCCGCT	1457
Db	1871	TGGCGCTCTGGGTAGCTCTGCTGTGGGACAGCGCCCAATTGCGCAGTGAAGCTGCCGA	1930
Qy	1458	GTCGCCCGGTCTTCTCCCAACGAGACCGGACGGCATCAAGAAATGCGAGTGTCAAGTGTG	1517
Db	1931	GTCGCGAATGACCCCAAGGTCTGGGCGAAGTG-----GAGCGATGAGTGTGACCAATTGG	1981
Qy	1518	CTTGCAATGGCGAGGTGAGACAGGTATCTACACAGTGTGAGACATGTGCTGTGGCAGAG	1577
Db	1982	CTATGAAACAGCGGATGAGACAGGTATCTACACATGTGGCCACATGTGCTGTGTAAGC	2041
Qy	1578	CTGCGGCTGGCGCTCAAGCGACAGAGCCCGGCGCTGTGGCCCATCTGCGCGGCGCCAT	1637
Db	2042	CTGTGGCTGGCGCTCAAGAAAGGCTGTGACGCGCTGTGGCCCATCTGCGCGGCGCCAT	2101
Qy	1638	CAGGAGCTCAATTAAGATCTACAG	1661
Db	2102	CAGGACATCATCAAGACCTTACCG	2125

RESULT 8			
LOCUS	U87864	2207 bp	mRNA linear
DEFINITION	Homo sapiens neutralized homolog (h-neutralized) mRNA, complete cds.		
ACCESSION	U87864		
VERSION	U87864.1	GI:315790	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBLISHED
1 (bases 1 to 2207)	Mammalia, Eutheria, Primates, Catarrhini, Homiidae, Homo.				2 (bases 1 to 2207)	Yoshida, H., Tsuniki, H., Ito, K., Ueno, M., Nakao, M., Oka, K., Tada, M., Kochi, M., Kuretsu, J., Ueno, Y. and Sawa, H.	Identification of a human homolog of the Drosophila neutralized gene within the 10q25.1 malignant astrocytoma deletion region	Oncogene 16 (8), 1009-1019 (1998)	
9819875					9819875				
2 (bases 1 to 2207)	Yoshida, M., Nakamura, H., Tsuniki, H., Makino, K. and Sawa, H.	Direct Submission	Submitted (30-JUN-1997)		2 (bases 1 to 2207)	Yoshida, M., Nakamura, H., Tsuniki, H., Makino, K. and Sawa, H.	Direct Submission	Submitted (30-JUN-1997)	
University School of Medicine, Honjo 2-2-1, Kumamoto 860, Japan					University School of Medicine, Honjo 2-2-1, Kumamoto 860, Japan				
Location/Qualifiers					Location/Qualifiers				
1. .2207	/organism="Homo sapiens"				1. .2207	/organism="Homo sapiens"			
/mol_type="mRNA"					/mol_type="mRNA"				
/db_xref="taxon:9606"					/db_xref="taxon:9606"				
/tissue_type="brain"					/tissue_type="brain"				
/dev_stage="fetal"					/dev_stage="fetal"				
1. .2207	/gene="h-neutralized"				1. .2207	/gene="h-neutralized"			
411. .2135	/gene="h-neutralized"				411. .2135	/gene="h-neutralized"			
/note="ring finger protein, similar to Drosophila neutralized, encoded by Genbank Accession Number L12218"					/note="ring finger protein, similar to Drosophila neutralized, encoded by Genbank Accession Number L12218"				
/codon_start=1	/product="neutralized homolog"				/codon_start=1	/product="neutralized homolog"			
/protein_id="AAC17474.1"	/db_xref="GI:3157991"				/protein_id="AAC17474.1"	/db_xref="GI:3157991"			
/translation="MGNPPSIPSLPRGNBPARGRHNPDKSISGPPPTSHRCHN					/translation="MGNPPSIPSLPRGNBPARGRHNPDKSISGPPPTSHRCHN				
KQKRSVLPSSGGLPATPPLPHPTKSSQILMDISHKAVKQASFCAITPTSNPVL					KQKRSVLPSSGGLPATPPLPHPTKSSQILMDISHKAVKQASFCAITPTSNPVL				
YVQRLKITTKQCCSGALRGFTSKQPSRIHPDSLPKACPDIVSOSGFPAKLPS					YVQRLKITTKQCCSGALRGFTSKQPSRIHPDSLPKACPDIVSOSGFPAKLPS				
PANSGNIIAFVPSDKKGVFPHINDSAVMLPSSGVTADPLMALVDVGLTRGVLDLS					PANSGNIIAFVPSDKKGVFPHINDSAVMLPSSGVTADPLMALVDVGLTRGVLDLS				
ELVPCDLRPFVPSFALRPSLRADADPARLSVLCIDLVPGADDEAPAGCPIPONK					ELVPCDLRPFVPSFALRPSLRADADPARLSVLCIDLVPGADDEAPAGCPIPONK				
SINSGHRLPAOLGDLRFLALGAHVRILDSQTVARVHGSRDEALVTSRPRV					SINSGHRLPAOLGDLRFLALGAHVRILDSQTVARVHGSRDEALVTSRPRV				
ASRTFVATYSGGARPRALSIGRTTCPTLRAPDLPPSRVALYDRSEFAVCPVPG					ASRTFVATYSGGARPRALSIGRTTCPTLRAPDLPPSRVALYDRSEFAVCPVPG				
LHSDIILGLVNNADGHLSTNNGAAGQGLCVDSQPMPLFGHLGTTQRIIGASTT					LHSDIILGLVNNADGHLSTNNGAAGQGLCVDSQPMPLFGHLGTTQRIIGASTT				
LAEGGIPSLPCSPSPSTPSPLGSRSLSDPLSTCSSGPISSAGTAPNSPVSLPES					LAEGGIPSLPCSPSPSTPSPLGSRSLSDPLSTCSSGPISSAGTAPNSPVSLPES				
PVTGLQMSDECTICYEHAVDVITYTCGHMCLCYACGLRLKALHACCPICRRPIKD					PVTGLQMSDECTICYEHAVDVITYTCGHMCLCYACGLRLKALHACCPICRRPIKD				
IIKTYRS"					IIKTYRS"				
1971. .2090	/gene="h-neutralized"				1971. .2090	/gene="h-neutralized"			
/note="ring finger motif"					/note="ring finger motif"				
872 c 673 g 322 c					872 c 673 g 322 c				
340 a					340 a				
872 c 673 g 322 c					872 c 673 g 322 c				
24.4%	Score 409.2;	DB 9;	Length 2207;		24.4%	Score 409.2;	DB 9;	Length 2207;	
56.6%	Pid. No. 2.7e-35;				56.6%	Pid. No. 2.7e-35;			
0;	Mismatches 628;	Indels 60;	Gaps 5		0;	Mismatches 628;	Indels 60;	Gaps 5	
Query Match					Query Match				
Beat Local Similarity					Beat Local Similarity				
Matches 896;	Conservative				Matches 896;	Conservative			
99	GATCTCTGGGAGAGGCGCGCTTCCAGCGCGCAAGGCAAGAGCAAGAGCTGCGGCTGGA	158			99	GATCTCTGGGAGAGGCGCGCTTCCAGCGCGCAAGGCAAGAGCAAGAGCTGCGGCTGGA	158		
581	GCTCCCAAGCAGCGCGCTCTTCCACCGGCAACCAAGGCTCCCAAGATCTCATGGA	640			581	GCTCCCAAGCAGCGCGCTCTTCCACCGGCAACCAAGGCTCCCAAGATCTCATGGA	640		
159	CGGCACTCGCGCGCGGCGCACACGCGCGCAAGCTTCTGCAATGGCGTCAAGTTCAAGCA	218			159	CGGCACTCGCGCGCGGCGCACACGCGCGCAAGCTTCTGCAATGGCGTCAAGTTCAAGCA			

279 GAGCGCGCGCTTGGCTTACCGCGACGATCCGTGCTCATAGAGGCCAGGA 338
 412 GAGCGCGCGCTTGGCTTACCGCGACGATCCGTGCTCATAGAGGCCAGGA 471
 339 CATCCCAAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 398
 472 GCTGCCAAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 531
 399 GCCCGAAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 458
 532 GCTCGAAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 591
 459 GCTGTTTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 518
 592 TGTCTTCAACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 651
 519 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 578
 652 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 711
 579 GAGCGCGCTTGGCTTACCGCGACGATCCGTGCTCATAGAGGCCAGGA 638
 712 TAGCGAGCTGG-----TGTCCCGAGCTGTGCGCGCGCGCTTTCAGCGCCT 762
 639 CTGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 698
 763 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 811
 699 GGTGTTGCGCGAGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 758
 812 -----AGCTATGCGAGCTTCAACGCTGCGCGCGCGCTACGCGCGCGCT 861
 759 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 818
 862 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 921
 819 GCGCGCTTGGCTTACCGCGACGATCCGTGCTCATAGAGGCCAGGA 875
 922 GCGCGAGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 981
 876 GTCGCGCGAGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 935
 962 CCTCGAAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1041
 936 CTGCGAGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 995
 1042 CACGAGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1101
 996 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1055
 1102 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1161
 1056 CAACGAGTACGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1115
 1162 GCGCGAGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1221
 1116 GCGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1175
 1222 CCGCGTGGCGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1281
 1176 GCGAGTCTCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1235
 1282 GAGGCTGACCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1341
 1236 GAGGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1295
 1342 GAGGCGCGCTGCGCGGACCTGTCACGCGCGCGCTACGCGCGCGCT 1401
 1296 CGGTACCTGCGAGTCCAGCGCGCTGTCACGCGCGCGCTACGCGCGCGCT 1348
 1402 CTCGACTATCTGCGCGGAGCGCGCTGTCACGCGCGCGCTACGCGCGCGCT 1461
 1349 -----GCTCCGAGAGCGATGATTCAGATATGACCTTCAGTGTCAACAGTTC 1397

Db 1462 AACCTGCCAGTGCCTTGGCGAGCGCGCTGTCAGACCCCTTGTCTCAGACGTGAGCTTC 1521
 GY 1398 CTCCTCGGAGTATGATGATTCCTGTCAGCGCGCGCGCTACGCGCGCGCT 1457
 Db 1522 TGGCGCTTGTGATGCTGTGTCAGCGCGCGCGCTACGCGCGCGCT 1581
 GY 1458 GTCCCGGTGTTCTCCCGAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1517
 Db 1582 GTCCCGGTGTTCTCCCGAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1632
 GY 1518 CTTCGAGTGGAGGTCAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1577
 Db 1633 CTATGAGACCGCGTGGAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1692
 GY 1578 CTGCGCGCTGCGGTCAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1637
 Db 1693 CTGTCGCTGCGGTCAGCGCGGACCTGTCAGCGCGCGCTACGCGCGCGCT 1752
 GY 1638 CAAGAGCTGATTAAGTCTTACAG 1661
 Db 1753 CAAGAGCTGATTAAGTCTTACAG 1776

RESULT 10
 BC026336 4011 bp mRNA linear PRI 08-APR-2002
 LOCUS Homo sapiens, neuralized-1 like (Drosophila), clone MGC:26482
 DEFINITION IMAGE:4812302, mRNA, complete cdf.
 ACCESSION BC026336
 VERSION BC026336.1 GI:20070954
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 4011)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shireki
 Tothlyuk and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IPAX Plate: 32 Row: 3 Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4758799.
 FEATURES
 source
 1..4011
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="locusid:9148"
 /db_xref="taxon:9606"
 /clone="MGC:26482 IMAGE:4812302"
 /tissue_type="Brain, hippocampus"
 /clone_id="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

CDS

98...1822
/codon_start=1
/product="neuralized-like (Drosophila)"
/protein_id="AAH26336.1"
/db_xref="GI:20070955"
/translation="MGNNPSSIPSLPRGNPSRAPRPHONLKDISGPPVTSRCH
KQKCGAVLPSGLPAAPLLFPHPTKQSOILMDISKAVROASPCNALITFSRPLVLI
YEOVRLKTKKQCCNMGALRLGTSKDPISIHDSIPRYACPLVVSQSFMAALBE
FANEONIIFWVDRKGRVFRINDSAVMLEFSGVTRADPLMALVDVGLRQVLDLS
BLVLPDLRPSFTALRPSLRREADARLSVLCDLNPGADGDEAAPASCPIDN
SINSHORALPAOLDLRFHALPAGAHYIIDEQTVARVHEDRDLAVTSRPARV
AETIPKVTSGSGARPGALPSGVTCDPGLPADLPSPBALVDKBRFARVPG
LHSGDILGLVYNAADGELHSHNGAAGMOLCNDASQPLMLRLHGTGQIRLIGSTI
LAEGLIPSLPSCPASTPSPALSGSLDPLSTGSSGPGSSAGTANSPISLES
PVTGAGWSDECTICYEHAVDVTIYTCGMELCYAGLKLKALHACPTCKRPYK
IKTVRS"

BASE COUNT 667 a 1303 c 1208 g 833 t

Query Match 24.3%; Score 407.6; DB 9; Length 4011;
Best Local Similarity 56.5%; Pred. No. 3.3e-35;
Matches 895; Conservative 0; Mismatches 629; Indels 60; Gaps 5;

QY 99 GGTCTGGGGGAGAGCGCGCTTCCACGCGAGGCGCAAGAGCGTGGGCTGA 158
DB 268 GCTCCAGCGCGCGCGCTGCTTCCACCGCACACCAAGAGGCTCCAGATCTCATGSA 327
QY 159 CGGCACTCGCGCGCGCGCACAGCGCGCAAGCTTCTGCATGAGCTTACGCA 218
DB 328 CCTCAGCGCAAGAGCTGTCAAGAGGAGCGAGCTTCTGCAGCGCTTACGCA 387
QY 219 GCGGCGCATCGGCTGTACAGAGAGTGGCGCTGCGCTGAGCGCGCGCTGCTG 278
DB 388 CGCGCGGCTCTCATCTACAGAGAGTCAAGCTGAGATCAAGAGAGAGTGTGCTG 447
QY 279 GAGCGGCGGCGGCTTGGCTTCAACGCGCGCAAGTCCGCTGCTGAGCGCGCGCA 338
DB 448 GAGCGGCGGCGGCTTGGCTTCAACGCGCGCAAGTCCGCTGCTGAGCGCGCGCA 507
QY 339 CATCCCAAGTACGCTGCGCGCGCAAGCTGCTGAGCGCGCGCTTACGCGCGCAAGGCACT 398
DB 508 GCTGCCAAGTACGCTGCGCGCGCGCAAGCTGCTGAGCGCGCGCTTACGCGCGCAAGGCACT 567
QY 399 GCGCGAGAACCTGCGCGCTGCGCGCAAGCTGCTGAGCGCGCGCTTACGCGCGCAAGGCACT 458
DB 568 GCGTGAAGATTTGCAATGAGGCAAGTCAATGCTGCTGAGCGCGCGCTTACGCGCGCAAGGCACT 627
QY 459 CGGTCTACAGGCTGAGAGAGCGCGCGAGCGGCTTCTTCAAGCGCGCGCGCTGAGCGCGCTG 518
DB 628 TGTCTTCCAGCGATCAAGAGTCTGCTGCTGCTGCTGCTTCAAGCGCGCGCGCTGAGCGCGCTG 687
QY 519 CGGCGCGCTTGGGCGCTCATTTGATGTCTACGCGATCAAGAGTGTGAGCTTCTGAG 578
DB 688 GAGCGCGCTTGGGCGCTTGGTGAAGCTTACGCGCTCAAGCGCGCGCTTCAAGCGCTTGA 747
QY 579 GAGCGCGCTTGGTGAAGCTTACGCGCGCGCGCTTCAAGCGCGCGCTTCAAGCGCGCTG 638
DB 748 TAGCGAGCTG-----TGTCCCGAGCTGTCTGCGCGCGCTTCTTCAAGCGCGCTT 798
QY 639 CTGCGCGCGCGAGAGCGAGCGAGCGCACTTCAAGAGAGCTGAGAGAGAGCA 698
DB 799 GCGCGCGCGCTGCTGCGCGCGAGAGCGAGCGAGCGCGCTTCTGCTG-----847
QY 699 GGTGTGTGCGCAAGCTGCGCGAGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCA 758
DB 848 -----AGCTTATGCGAGCTCAAGCTGCGCGCGCGAGCGAGCGAGCGCGCGCGCG 897
QY 759 GCGCGCGCGCGCGCGCATTTCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCTG 818
DB 898 GCG 957
QY 819 GCGCGCTACGAGCGCGAGCTTCAAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 875

DB 958 GCGCGAGCTGAGAGCGAGCTGCTTCCAGCGCGCGCGCGCGCGCGCGCGAGCTCCGCA 1017
QY 876 GTCCGCGGAGCGCGCAAGTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 935
DB 1018 CCTGAGCGAGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1077
QY 936 CTCCGAGCGCGCGCTGCT 995
DB 1078 CACGAGCT 1137
QY 996 GCG 1055
DB 1138 CG 1197
QY 1056 CAGAGAGCTGCG 1115
DB 1198 GCGCGAGCTGCG 1257
QY 1116 GCG 1175
DB 1258 CG 1317
QY 1176 CGAGCTGCTCTGAGGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
DB 1318 CGAGCTGAGCTGAGGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1236 GCGAGGCTGCTGAGGATCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
DB 1378 GCGAGCGCTTGTGATCTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
QY 1296 CGGTACCGTGAAGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1348
DB 1438 CTCACTATCTGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
QY 1349 -----GCTCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1397
DB 1498 AACCTGCGCGAGCG 1557
QY 1398 CTCTGCGGATCTGAGTCACTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1457
DB 1558 TGCGCTCTGAGGATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1617
QY 1458 GTCGCCGAGTCTTCCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
DB 1618 GTCCGAGTACCGCGAGTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1668
QY 1518 CTTCAGTGGAGAGTGAAGAGCTGATCTACAGCTGTGAACAGATGCTGTGCGCAAG 1577
DB 1669 CTATGAACAGCGGTGAGAGCGGTCACTTACATGATGAGCAATGCTGTGCTACGC 1728
QY 1578 CTGCGCGCTGCGGCTCAAGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1637
DB 1729 CTGTGCGCTGCGGCTCAAGAGAGCTGCTGAGCGCGCGCGCGCGCGCGCGCGCG 1788
QY 1638 CAGAGAGTCAATTAAGTACAG 1661
DB 1789 CAGAGAGTCAATTAAGTACAG 1812

RESULT 11
MNEURALI 2445 bp mRNA linear ROD 07-JAN-2000
LOCUS MNEURALI 2445 bp mRNA linear ROD 07-JAN-2000
DEFINITION Mus musculus mRNA homologous to protein encoded by Drosophila
neutralized gene.
ACCESSION Y15160
VERSION Y15160.1 GI:6688678
KEYWORDS neutralized gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Pavlopoulos, E., Prinos, P., Kiplpatrick, M., Tsipouras, P. and

Maschonas, N.K.
Unpublished
Journal
Reference

AUTHORS Moschonas, N. K.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-1997) N.K. Moschonas, Dept of Biology, University of Crete and IMBB-FORTH, P.O. Box 1527, Heraklion 711 10, Crete, Greece

REMARK Revised by (3)
REFERENCE 3 (bases 1 to 2445)
AUTHORS Moschonas, N. K.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) N. K. Moschonas, Dept of Biology, University

REMARK	Revised by author 23-MAR-1998
FEATURES	Location/Qualifiers
SOURCE	1..2445

CDS

BASE COUNT	435 a	843 c	700 g	467 t
ORIGIN				

Query Match	23.9%	Score 399.6;	DB 10;	Length 2445;
Best Local Similarity	56.2%	Pred. No. 2.8e-34;		
Matches 890;	Conservative 0;	Mismatches 634;	Indels 60;	Gaps 5;

QY	99	GGTCTTGGGCGAAGGCGCGCCCTTTCAGCGCCAGGCGCAAGGCAAGAACTGCGGCTGGA	158
Db	171	GCTCCCGGCGACGCGCGCTGCTCTTTCACCCCGCACTAAGGGCTCCCGAGATCTCATGGA	230
QY	159	CGGCACTCGCGCGCGGCGCACAGCGCGCAAGCTTCTGCAATGAGCGTCAAGTTACGCA	218
Db	231	CCTCAGCCACAAGGCGCGCTTCAGAGGCGAGCCAGCTTCTGCAATGCAATCACTTCAAGTA	290
QY	219	GCGGCGCATTCGGCGCTGACGAGCAGGAGGCGGCTGCGCGCTGAGCGCGCTGCGCTG	278
Db	291	CCGCGCGGTGCTACTCTACGAGCAAGTCAGGCGTGAAGATCAACAAGAAAGCATGTCGTCG	350
QY	279	GAGGCGGCGCGCTGCGCTTCGCGCTTCAACCGCGCACATTCGTCGCTCATGAGCGCCACAGA	338
Db	351	GAGCGGGGCGCTCGCGCACTTGGCTTCCACGAGCAAGGACCTTCCCGCATTCACCCCGATTC	410
QY	339	CATCCCGCAAGTAACGCGCTGCGCGCGACCTGATCAACGCGCGCGCGCGCTCATGAGGCGAAGGCACT	398
Db	411	GCTGCGCCAAAGTACGCTGCGCTTCCTGACCTGATCTGATGTCTTGAAGTGGCTTCTGGCGCAAGCATTT	470
QY	399	GCCCGAGAACCTGGCGCTGCGCGCACACAGGTGCTGGCTCACTTGGCGCGACCGCGCACGCGCG	458
Db	471	GCTGAGGAGATTGGCCAAACGAGGGGAAACATATTGCTCTTGGGTGGAGCAAGAAAGGCGCG	530
QY	459	CGTGTCTTACAGCGTGAACGACGAGCGAGCGGCTCTTTCACATGCGGCGTGGCCGTGGG	518

D	531	CGTCTTCTACGGATGAAATGAGTCAGGTGTATGCTTTTCTTCAATGAGGGATCCGGACGGC	590
Q	519	CGGCGCGCTGTGGGCGCTAATTGATGTCTACGGGATTCACGACGAGGTGAGCTTCTTGA	578
D	591	GGACCGGCTCTGGGCGCTGGTGGACGTTACCGGCTCCACCGGGGTGTCCAGCTGTACGA	650
Q	579	GAGGCGCTTGGCTGACACGGTGAACGGCCGGCGGCTCAGGCAAGGCCGCTTACGGGCTG	638
D	651	CAGCGAGCTGG-----TGGTGGCCGACTGGCGGCGGCTTCTTACCGGCGCT	701
Q	639	CTTGGCGGCCGACGACGACGACCGCGGCAACTTTCAGCAACAGAGCTTCGAGAACAAACA	698
D	702	GCGGCGGCGCTCGCTGGGTGGAGGGG-----GATTAGGGGGCG	740
Q	699	GATGTGTGGCCAGCTGGGGCACTTGGGCTTGGGCGGCGCCCGGGGCCACCGGACGGCA	758
D	741	CTGTGCGGTGAGCTGTGGCGAAGCTTCAACGTGGCGGGGAGCGACGACGACGACGCGAC	800
Q	759	CGCGCGGGGCGCGCGCAATTCCGTGGGGGGCCCGTAGAGCGCCCGGCGCGCTGTGGCC	818
D	801	GCTTGGCGGGTGTGGCCCGATTCCCGGAGAACTGGCTCAATTTCAGAGACAGCGCGGCTTGC	860
Q	819	GGGCGCTACTTGAAGGCGGACCTTGGGCTTTCACGCGCAACGCGGGGCGG---ACGTGAGCT	875
D	861	GGGCGAGCTGAGCGGCGACCTTGGGCTTTCACGCGCTTTCGCGCGCGCGGACGTCGGCAT	920
Q	876	GTCGGGCGAACCGGCAAAATGTGGCTTGGCGACCGCGGGGCCGACGCGGAGCTGTCTT	935
D	921	CTTGAAGAGACAGACGGTGGCGGCGCGGTGGAGGACACGGGGGGGAGCGAGGCGGCTGTCTT	980
Q	936	CTTCCGAGCGCCCGCTGGCGGCGCGCGGACGAGCGCTTCTGTGAGAGGTGGCGGTCCGGAGCT	995
D	981	CACACAGCGCGGCTGTGGCGGTGGCGCGAGAACATCTTATCAAGGTTCACCGGCTTCCGGTGG	1040
Q	996	GGGCGGCGCGGGCGGCGGTGGCGCTTGGGCACTCAAGTGGTGGACCCCGGGCGGTGCTACGGCC	1055
D	1041	CGGCGGCGCGGGCGGCTGTGCTGTCCGGGGTCCACAGGTGTGACCCCGGACGCTTGGCGC	1100
Q	1056	CAACGAGCTGACCGGCGGACGACGACGACGCGCTTGGACCGGCAAAAGATACTGGGTGTGGAC	1115
D	1101	CGGAGACCTTGGCGCTTTCAGCGCCCGAGGCGCTGTGTGAACCGGAAAGATTTGGGGGAGTGG	1166
Q	1116	GCGGCGCGGGGCGGCTGGCGGAGCGGCGGCGGACGCGCTCAGCTTCAACGCTGCGGGCGCGG	1175
D	1161	TGGGTGTCCGGGCGCTTGTGACAGCGGCGGACATCTTGGGCTGTGTGTCAACGCGGACGG	1220
Q	1176	CGACGTGTCTTGGGCGATCAACCGGGCGTCCGCGGGCGGCGGCTGTGTGGGTGACACACAC	1235
D	1221	AGAGCTGACACTTGAATGACAAACGCGCGGGCGGCGGGCAATGAGTGTGGCGGTGATGCTTC	1280
Q	1236	GCAAGCGCTGTGGGCGCTTCTTGGCGGTGGCGGCGGCGGAGTGTGGCGGCTCTCTCT	1295
D	1281	GCAAGCGCTGTGGATGTCTTTCAGGCTGTGATGGGCGCATACGAGAGGTCCGACATCTCTGG	1340
Q	1296	CGGTACCTTGAAGTTCAGCGCCCTTGGACACGAGCTCATCAAGGATCTCTCAAGG-----	1348
D	1341	CTTCAACATATGACTGAACGGGGGTGGCCCATCTCTCCCTTGCCTCACTTGCCTCACTCC	1400
Q	1349	-----GCTCCAGGACGATATGTGATTCAGATATATGACTTTCAGTGTTCACAGTCC	1397
D	1401	AACCTCAACCAATGAGCCCTTGGGCAATCGGCTCTTGAACCCCTGTCTAGACACTTGGGTTCC	1466
Q	1398	CTTCTGGGCAATCTGAGTTCATCCCTTGGTGAAGAGGCGCCGAGCTCCCGGCTGAGGCGCGGT	1455
D	1461	TGGGCGCTTGAATGTGTCTTGGAGGAGGACAGGCCCGCACTCACTTGTAGGCGCTGGCCGA	1520
Q	1458	GTCGCCCGGTGTCTCCCGACCGAGGCGGCGAGGAGCATCAAGATAGCGAGTGCACGGTGTG	1517
D	1521	GTCAACGGTGAACCCAGATCTTGGGCGAGTG-----GAGTATGAAATGCACCATTTGG	1571
Q	1518	CTTGAATGGAGGTGTGACACGGTCACTTACGGTGTGACACATGTGGCTGTGGCCAG	1577

QY 1296 CGGTACCTGCAAGTCCAGCCCTGCGACAGACTCCATCAGGGTCCCTCAGCG----- 1348
 DB 1559 CTCACACATCATGATGAGAACGGGGGTGGCCCATCTCCCTGCTGACCTGGCCCACTCC 1618
 QY 1349 -----GCTCCAGAGAGATATGTATTCATGATATGACCTTGATGTCAACCATGTC 1397
 DB 1619 AACCTCACACAGAGCCCTGAGGATCCGCTCTCTGACCCCTGCTCAGACACTGCGGTTTC 1678
 QY 1398 CTCCTGCGCATCTGATGATCATCCCTGAGACGGCCCGCCAGCTCCCGCTGAGCCCGCGGT 1457
 DB 1679 TGGGCTCCCTAGGCTGCTCTGCTGAGAGAGACGCCCACTGACCTGATAGCTGCTCCCA 1738
 QY 1458 GTCCCCCGGTCTCTCCCAACCGGAGCCGAGGACATCAAGAAATGGCGAGTGAACGGTGTG 1517
 DB 1739 GTACCCGGTGAACCCCAAGGTCTGGGGCAGTG-----GAGTGAATGAATGACCATTTGG 1789
 QY 1518 CTTTCATGGCGAGGTGAGACACGGTCACTTACAGTGTGACACATGTGCTGCTGCGACAG 1577
 DB 1790 CTATGAAACAGCAGTGTGACACAGTCACTACAGTGTGACACATGTGCTGCTGCTACTTC 1849
 QY 1578 CTGCGGCTGCGGCTGACAGGACCGGCGCTGCTGCGCCATCTGCGCGGCGGCGCAT 1637
 DB 1850 CTGTGCTGCGGCTGCTGACAGGCGCTGACCGCTGCTGCGCCATCTGCGCGGCGGCGCAT 1909
 QY 1638 CAAGACGTCATTAAGATCTACAG 1661
 DB 1910 CAAGGACATCATCAAGACCTACCG 1933

RESULT 13

AF401228 3630 bp mRNA linear ROD 15-OCT-2001
 LOCUS Mus musculus neutralized (Neur1) mRNA, complete cds.
 DEFINITION AF401228
 ACCESSION AF401228.1 GI:15420882
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS 1 (bases 1 to 3630)
 TITLE Vollrath, B., Pudney, J., Aas, S., Leder, P. and Fitzgeraid, K.
 JOURNAL Isolation of a murine homologue of the *Drosophila* neutralized gene, a gene required for axonemal integrity in spermatozoa and terminal maturation of the mammary gland
 Mol. Cell. Biol. 21 (21), 7481-7494 (2001)

JOURNAL
 MEDLINE 21470324
 PUBMED 11585928
 REFERENCE 2 (bases 1 to 3630)
 AUTHORS Vollrath, B., Pudney, J., Aas, S., Leder, P. and Fitzgeraid, K.
 TITLE Direct Submision
 JOURNAL Submitted (20-JUL-2001) Department of Genetics, Harvard Medical School, 200 Longwood Ave, Boston, MA 02115, USA
 FEATURES
 source
 1..3630
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..3630
 /gene="Neur1"
 156..1829
 /note="similar to *Drosophila* melangaster neutralized"
 /codon_start=1
 /product="neutralized"
 /protein_id="AAK97495.1"
 /db_xref="GI:15420883"

gene
CDS

/translation="MGQITRTNTHDSIGSPVPSHRCHNKHCPPTLSGGGUPAT
 PLTFPHTKGSOILMDLSHAKVROQSPFNATTFSPRPVLIYQYRLKTKKCCMSG
 ALRGTGSKDPSRHPDLSKYACPDVLSQSGFMALKEBPAANGNIIAFVDDKGR
 VVPIRINSAAMT.PFSGRTVDPDLMALVNYGLTRGQVLLDSLVN.PDCIRPSPALR
 RSLRCEADEARLVSICDLPNPGADGDDGAPPGCPITQNSLNSOHSALEPAQIDGD
 LRFHARAGAHVRIIDQTVARLEHGRDERALVFTSRPVSVAETIPIKTRSGGGRAG

BASE COUNT 667 A 1119 C 1082 G 762 T
 ORIGIN

Query Match 23.5%; Score 393.2; DB 10; Length 3630;
 Best Local Similarity 55.9%; Pred. No. 1.2e-33;
 Matches 886; Conservative 0; Mismatches 638; Indels 60; Gaps 5;

ASBPGVTCDBETRLPADLPSPBALVDNRKFWAVRPGPLHSGDILGLVYVNDGL
 HLSHNGAAGMQLCVDSQPLMPLPSLHGAITQYRIIGSTIINTERGSGSLPSPASTP
 TSPBALIGTSGDPLGDSAGPLAPNSBPSVLPBSPVPTGACOMDBECTICY
 RHAVDTVYTCGHMCLCYSGSLRLKXLAACCPICRRIRIINDIINTYRS"
 QY 99 GGTCTGAGCGAGGCGCGCGCTTCCACGCGAGGCGCAAGGCAAGAACTGCGGCTGGA 158
 DB 275 GCTCCCGGCAACGCGCGCTTCTTCCACCCCACTAAAGGCTCCAGATCTCATGGA 334
 QY 159 CGGCACTCGGCGCGGCGCAACCGCGGCAAGAGCTTCTGACATGGGTCAGTTACGCA 218
 DB 335 CTTGAGCCACAAAGCGCTTCAGAGAGGCAAGGCGGCTTCTGACATGCACTTACGTA 394
 QY 219 GCGGCGCATCCGCGCTGATGAGAGAGTCCGCGCTGCGCTGAGCGCTGAGCTG 278
 DB 395 CGGCGGAGTGTCACTGAGAGAGTCAAGGCTGAATGACCAAGAGGCAATGCTGCTG 454
 QY 279 GAGCGGCGGCTGCGCTTCCGCTTCAACGCGGCAAGATCCGTGCTCATAGCGCCGAG 338
 DB 455 GAGCGGCGGCTGCGCTTCCGCTTCAACGAGAGAGGCGCTTCCGCTTCAACCGGCTC 514
 QY 339 CATCCGCAAGTACGCTGCGGAGCTGAGTCAACGCGGCGGCTTCTGAGGCAAGGCT 398
 DB 515 GCTGCGCAAGTACGCTGCGGAGCTGAGTCAACGAGTGTCTGAGAGTGTCTTGGGCAAGGCT 574
 QY 339 GCGCGAGAACCTGCGGCTGAGGAGACCGGCTGAGTGTCTTCTGAGGCGGCGGCG 458
 DB 575 GCTGAGAGAGTGTGCAACGAGGAGCAATCATTTGCTTGGGAGTGTGCAAGAGGCGG 634
 QY 459 GGTGTTCTACAGCGTGAACGAGCGGAGCGGCTTCTTCACTGCGGCGTGGCGG 518
 DB 635 GGTCTTCTACAGCGTGAACGAGCGGAGCTGAGTGTCTTCTTCACTGAGGAGTGTGCGGCT 694
 QY 519 GCGGCGGCTGAGGCGCTCATTTGATGTGACGAGTCAACGAGGAGTGTGAGCTTGTGA 578
 DB 635 GAGCGGCTTGTGAGGCGCTTGTGAGAGTGTGAGGCTTGTGAGGAGTGTGAGGCTTGTGA 754
 QY 579 GAGCGGCTTGTGAGGCGCTTGTGAGGCGGCTTGTGAGGAGGCGGCTTGTGAGGCGG 638
 DB 755 CAGGAGGCTG-----TGCTGCCGAGCTGCGGCGGCGGCTTGTGAGGCGGCT 805
 QY 639 CTTGCGGCGGCAAGGAGCGGCGGCAAGGCGGCGGCAAGGAGTGTGAGGAGCAACCA 698
 DB 806 GCGGCGGCGGCTGCGGCTGAGTGTGAGGCGG-----GATGAGGCGG 844
 QY 699 GGTGAGGCAAGGAGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 758
 DB 845 CTTGAGTGTGAGGCGGCTGAGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 904
 QY 759 GCGCGGCGGCGGCGGCTTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 818
 DB 905 GCTGCGGCGGCGGCGGCTTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 964
 QY 819 GCGGCTTGTGAGGCGGCGGCTTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
 DB 965 GCGGCGGCTGAGGCGGCGGCTTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1024
 QY 876 GTCGCGGAGCGGCAAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 935
 DB 1025 CTTGAGCAAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTCT 1084
 QY 936 CTTGAGGCGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 995
 DB 1085 CACGAGCGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1144
 QY 996 GCGGCGGCGGCGGCGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1055

Db 1145 GGGGCGAGCGGGCGGCTGTCCTTCGAGGATCACCACGATGATGACCTCGGACGCTGCGGCGC 1204
 Qy 1056 CAACGAGCTGCGGCGCCGACCCAGACGCGCTGCTGACCGCAAGATGATGAGTGGTGGGC 1115
 Db 1205 CGCGGACCTGCTTTCAGAGCCCGGAGGCTGTGTGACCGCAAGAGATTTCTGGGCGGTGTG 1264
 Qy 1116 GCGGCGGCGGCGCGTGCAGAGCGGCGGACGCGCTGACGCTTCAAGCGTGGCGGCGGCGG 1175
 Db 1265 TCGCGTCCCGCGGCGCTTTCGACAGAGGAGGACATCTTGGGCTGTGTGATCAAGCGGACG 1324
 Qy 1176 CGAGTGTCTCTTGGGCGATCAACGAGCGCTCGCGGCGCGCGCTGTGTGCTGACACAC 1235
 Db 1325 AGAGCTGCACCTGATGACACACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1384
 Qy 1236 GCAGGCGCTTGGGCGCTTCTTTCGCGCGTGCAGCGGCGGCGCTGCGGCGGCGGCGGCT 1295
 Db 1385 GCAGCGCTTGTGATGCTCTTTCAGCTGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 1444
 Qy 1296 CGGATCCCTGACATGACGCGCGCTGAGACGACGATCCATCAGAGTCCCTCAGCG----- 1348
 Db 1445 CTCACCATCATGATGATGACGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1504
 Qy 1349 -----GCTCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1397
 Db 1505 AACCTACCGACGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1564
 Qy 1398 CTCCTGCGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1457
 Db 1565 TGAGCGCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1624
 Qy 1458 GTCCCCCGTCTTCTCCCAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1517
 Db 1625 GTCAACGCTGACCCGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1675
 Qy 1518 CTTGATGAGCGAGGATGACAGCGATCTACAGCTGTGAGACATGATGCTGTGCGACAG 1577
 Db 1676 CTATGACACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1735
 Qy 1578 CTGCGGCGCTGCGGCTCAAGCGACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1637
 Db 1736 CTGTGCGCTGCGGCTCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1795
 Qy 1638 CAAGACGCTCATTAAGATCTACAG 1661
 Db 1796 CAAGACATCATTAAGATCTACAG 1819
 RESULT 14
 E26375 LOCUS 1641 bp DNA linear PAT 18-JUN-2001
 DEFINITION Neutralized protein, polynucleotide encoding said protein and
 E26375 antibody recognizing said protein.
 ACCESSION E26375.1 GI:13025071
 VERSION JP 1999137257-A/2.
 KEYWORDS unclassified
 SOURCE unclassified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1641)
 AUTHORS Motomi, N., Hideo, N., Mitsuhiro, Y. and Hideyuki, S.
 TITLE Neutralized protein, polynucleotide encoding said protein and
 JOURNAL antibody recognizing said protein
 PATENT: JP 1999137257-A 2 25-MAY-1999;
 SUMITOMO ELECTRIC IND LTD
 OS Unidentified
 PN JP 1999137257-A/2
 PD 25-MAY-1999
 PR 14-NOV-1997 JP 1997313211
 PC MOTOMI NAKADA, HIDEO NAKAMURA, MITSUHIRO YOSHIDA, HIDEYUKI SAKA
 PC C12N15/09, C07K14/47, C07K16/18, C12P21/02, C12Q1/68, G01N33/53//
 PC C12N15/09, C12R1:91, (C12P21/02, C12R1:19), C12N15/00, (C12N15/00, PC

C12R1:91)
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT Source 1..1641
 FT /organism="unclassified".
 FEATURES
 source
 1..1641
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 260 a 594 c 498 g 288 t 1 others
 ORIGIN
 Query Match 23.4%; Score 392.2; DB 6; Length 1641;
 Best Local Similarity 55.9%; Pred. No. 28-33;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;
 99 GATCTGGGCGAGCGCGGCTTTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 158
 Db 87 GCTCCGCGGCGAGCGCGGCTGCTTTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 146
 Qy 159 CGGCACTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 218
 Db 147 CTTGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 206
 Qy 219 GCGGCGCATCGGCTGATGACAGAGTGGCGGCTGCGGCTGTGTGCGGCGGCGGCGGCGG 278
 Db 207 CGGCGCGGCTCTCATCTACAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 266
 Qy 279 GAGCGGCGGCGGCGGCTTTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
 Db 267 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 326
 Qy 339 CATCCCGAGTACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 398
 Db 327 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 386
 Qy 399 GCGCGGAGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 458
 Db 387 GCGCGGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 446
 Qy 459 CGGATCTACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 518
 Db 447 GATCTTTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 506
 Qy 519 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 578
 Db 507 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 566
 Qy 579 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 638
 Db 567 CAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 617
 Qy 639 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 698
 Db 618 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 656
 Qy 699 GATGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 758
 Db 657 CCGTGGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 716
 Qy 759 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 818
 Db 717 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 776
 Qy 819 GCGCGTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
 Db 777 GCGCGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 836
 Qy 876 GTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 935
 Db 837 CCGTGGAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 896

PA (TIMM/) TIMMUSK T.
 XX Palm K, Timmusk T;
 PI
 XX WPI: 2003-174081/17.
 DR P-PSDB; ABUS6251.
 XX

Novel neutralized polypeptide, Neu useful for controlling cell proliferation and calcium signalling induced transcriptional processes and treating depression, pain, anxiety, cancer and neurodegenerative diseases -

Claim 6; Page 38-39; 75pp; English.

The invention relates to a purified neutralised (Neu) polypeptide comprising at least one neutralised homology repeat (NHR) domain and a C3Hc4 RING-zinc finger domain, and a polynucleotide encoding the polypeptide. The polynucleotide is useful for constructing a transformed host cell that expresses a Neu protein. Neu has neurogenic function and functions as a transcriptional regulator, as a calcium-signal transducer, in cell signalling and synaptogenesis, in memory learning, RNA processing, early immediate responses, death of specific cell populations, nuclear hormone receptor signalling and axonal path-finding. The Neu family of proteins presents a set of diagnostic and therapeutic targets, especially for treating cancer. Manipulating Neu expression and function is useful in controlling a variety of diseases including depression, pain, anxiety, neurodegenerative diseases and cancer. Sequences ABX75848-ABX75864 represent Neu polynucleotides of the invention.

Sequence 1675 BP; 231 A; 648 C; 561 G; 235 T; 0 other;

Query Match 100.0%; Score 1675; DB 25; Length 1675;
 Best Local Similarity 100.0%; Pred. No. 1.5e-257;
 Matches 1675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCAACACGGTGCACCGGACCCCTGACGACCCGAGCCCAACCGGCGGCTCTGAGCC 60
 DB 1 ATGGGCAACACGGTGCACCGGACCCCTGACGACCCGAGCCCAACCGGCGGCTCTGAGCC 60
 QY 61 ACCCGGCGTGTGCGGCGCCCGCGCCGAGCGACGCGCGCTCTGAGCGAGCGCGCC 120
 DB 61 ACCCGGCGTGTGCGGCGCCCGCGCCGAGCGACGCGCGCTCTGAGCGAGCGCGCC 120
 QY 121 TTCGACGCGGCAAGGCAAGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 121 TTCGACGCGGCAAGGCAAGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 CGGCGCAACAGCTTCTGCAATGAGCGCTCACTTCAACGACGCGGCGGCGGCGGCGGCGG 240
 DB 181 CGGCGCAACAGCTTCTGCAATGAGCGCTCACTTCAACGACGCGGCGGCGGCGGCGGCGG 240
 QY 241 CAGGTGCGGCTGCGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 DB 241 CAGGTGCGGCTGCGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 301 TTCACGCGGCAAGATCCGTCGCTCATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 301 TTCACGCGGCAAGATCCGTCGCTCATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 QY 361 GACCTGTGTCACGCGCGGCGGCTACTGAGCGCAAGGCACTGCGGAGAACTTGGCGCTGGC 420
 DB 361 GACCTGTGTCACGCGCGGCGGCTACTGAGCGCAAGGCACTGCGGAGAACTTGGCGCTGGC 420
 QY 421 GACGAGGTGCTGGCTTACTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 DB 421 GACGAGGTGCTGGCTTACTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

QY 481 GCGGAGCGGCTCTTCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 DB 481 GCGGAGCGGCTCTTCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 QY 541 GATGTCTAGGCACTCAACGAGAGGTGAGCTTCTGAGAGACGCTTCTGCTGACACGCTG 600
 DB 541 GATGTCTAGGCACTCAACGAGAGGTGAGCTTCTGAGAGACGCTTCTGCTGACACGCTG 600
 QY 601 ACGCGCGGCGGCTCAAGCAGAGCGGCTTCAAGCGGCTGAGCGGCGGCGGCGGCGGCGGCGG 660
 DB 601 ACGCGCGGCGGCTCAAGCAGAGCGGCTTCAAGCGGCTGAGCGGCGGCGGCGGCGGCGGCGG 660
 QY 661 GCGGCAACCTTCAACCAACGAGCTGAGAACCAACGAGTGTGAGCCAGCTGAGCGGCGG 720
 DB 661 GCGGCAACCTTCAACCAACGAGCTGAGAACCAACGAGTGTGAGCCAGCTGAGCGGCGG 720
 QY 721 CTGGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 DB 721 CTGGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 QY 781 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 DB 781 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 QY 841 CGCTTCAACGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 DB 841 CGCTTCAACGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 QY 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 DB 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 QY 961 GAGAGCGCTTCTGAGAGAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTC 1020
 DB 961 GAGAGCGCTTCTGAGAGAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTC 1020
 QY 1021 GGCATCAGCTGTGAGGACCGCGGCGGCTGTACCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 DB 1021 GGCATCAGCTGTGAGGACCGCGGCGGCTGTACCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 QY 1081 GCGCTGTGACCGGAAAGATGATCTGAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 DB 1081 GCGCTGTGACCGGAAAGATGATCTGAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 QY 1141 GCGGACGCGGCTCAGCTTCAACGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 DB 1141 GCGGACGCGGCTCAGCTTCAACGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 QY 1201 CGTCCGCGGCGGCGGCTGTGCGTGTGACCAACGAGCGGCTGTGAGCGCTTCTTTCGCG 1260
 DB 1201 CGTCCGCGGCGGCGGCTGTGCGTGTGACCAACGAGCGGCTGTGAGCGCTTCTTTCGCG 1260
 QY 1261 GTGCGCGGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
 DB 1261 GTGCGCGGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
 QY 1321 GTGCGCGGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
 DB 1321 GTGCGCGGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
 QY 1381 ACCAGGATCCATCAGAGGCTCTCAAGCGGCTCCAGAGAGATGATGATTAATGAGCC 1440
 DB 1381 ACCAGGATCCATCAGAGGCTCTCAAGCGGCTCCAGAGAGATGATGATTAATGAGCC 1440
 QY 1441 TTAGGTGTCAACAGATCTCTCTGAGCATGTAGTCAATCTGTGAGCGGCGGCGGCGGCGG 1500
 DB 1441 TTAGGTGTCAACAGATCTCTCTGAGCATGTAGTCAATCTGTGAGCGGCGGCGGCGGCGG 1500
 QY 1501 GCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
 DB 1501 GCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
 QY 1561 ATGTGCTGTGCAACAGCTGCGGCGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620

Db 934 GAGAGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 993
 Oy 1021 GGACATCACTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1080
 Db 994 GGACATCACTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1053
 Oy 1081 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1140
 Db 1054 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1113
 Oy 1141 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1200
 Db 1114 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1173
 Oy 1201 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1260
 Db 1174 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1233
 Oy 1261 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1320
 Db 1234 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1293
 Oy 1321 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1380
 Db 1294 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1353
 Oy 1381 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1440
 Db 1354 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1413
 Oy 1441 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1500
 Db 1414 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1473
 Oy 1501 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1560
 Db 1474 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1533
 Oy 1561 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1620
 Db 1534 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1593
 Oy 1621 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1688
 Db 1594 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 1641
 RESULT 3
 ABX75859 standard; cDNA; 1129 BP.
 ABX75859;
 31-MAR-2003 (first entry)
 Human Neu2 cDNA #2.
 Human; mouse; rat; gene; ss; neutralised polypeptide; Neu; NHR; cancer;
 neutralised homology repeat; CHHC4 RING-zinc finger domain; myogenesis;
 neurogenic function; transcriptional regulator; cell signalling; pain;
 calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 tumorigenesis; organ development; central nervous system; depression;
 neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 inductive tissue interaction; cytokine signalling; axonal path-finding;
 neurodegenerative diseases; cytoskeletal; antidepressant; analgesic;
 tranquiliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 Homo sapiens.
 US2002132293-A1.
 19-SEP-2002.

PF 14-MAR-2001; 2001US-0808387.
 XX 14-MAR-2001; 2001US-0808387.
 PR 14-MAR-2001; 2001US-0808387.
 XX 14-MAR-2001; 2001US-0808387.
 PA (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 XX Palm K, Timmusk T;
 PI Palm K, Timmusk T;
 DR WPI, 2003-174081/17.
 PT P-PSDB; ABUS6252.
 PT Novel neutralized polypeptide, Neu useful for controlling cell
 proliferation and calcium signalling induced transcriptional processes
 and treating depression, pain, anxiety, cancer and neurodegenerative
 diseases -
 PS Claim 6, Page 41, 75pp; English.
 CC The invention relates to a purified neutralised (Neu) polypeptide
 comprising at least one neutralised homology repeat (NHR) domain and a
 CHHC4 RING-zinc finger domain, and a polynucleotide encoding the
 polypeptide. The polynucleotide is useful for constructing a transformed
 host cell that expresses a Neu protein. Neu has neurogenic function and
 functions as a transcriptional regulator, as a calcium-signal transducer,
 in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.
 XX
 XX
 S0 Sequence 1129 BP; 158 A; 440 C; 372 G; 159 T; 0 other;
 Query Match 65.9%; Score 1103.6; DB 25; Length 1129;
 Best Local Similarity 98.8%; Pred. No. 8.2e-167;
 Matches 1112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 Oy 550 GGACATCACTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 609
 Db 4 GGACATCACTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 63
 Oy 610 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 669
 Db 64 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 123
 Oy 670 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 729
 Db 124 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 183
 Oy 730 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 789
 Db 184 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 243
 Oy 790 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 849
 Db 244 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 303
 Oy 850 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 909
 Db 304 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 363
 Oy 910 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 969
 Db 364 GCGCTGTGTGGAAGTGGAGCGCCCGGCGCTGGCGCGCCGCACTGTGCTTC 423

QY 970 TTCTGAGAGTGGGCGCTCCGGGGCTGGCGGGCGGGCGGGCTTGGGATCAGC 1029
 DB 424 TTCTGAGAGTGGGCGCTCCGGGGCTGGCGGGCGGGCGGGCTTGGGATCAGC 483
 QY 1030 TTCTGAGAGTGGGCGCTCCGGGGCTGGCGGGCGGGCGGGCTTGGGATCAGC 1089
 DB 484 TTCTGAGAGTGGGCGCTCCGGGGCTGGCGGGCGGGCGGGCTTGGGATCAGC 543
 QY 1090 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1149
 DB 544 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603
 QY 1150 CTGAGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1209
 DB 604 CTGAGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 663
 QY 1210 GGGCGCGCTGCTGGGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1269
 DB 664 GGGCGCGCTGCTGGGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 723
 QY 1270 GGGCG 1329
 DB 724 GGGCG 783
 QY 1330 CCATCAGAGTCTCTCAGCGCTCCCGAGAGCAGATGATGATGATGATGATGATGAT 1389
 DB 784 CCATCAGAGTCTCTCAGCGCTCCCGAGAGCAGATGATGATGATGATGATGATGAT 843
 QY 1390 AACAGTCTCTCTGAGCATCTGAGCATCTGAGCATCTGAGCATCTGAGCATCTGAGCAT 1449
 DB 844 AACAGTCTCTCTGAGCATCTGAGCATCTGAGCATCTGAGCATCTGAGCATCTGAGCAT 903
 QY 1450 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509
 DB 904 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963
 QY 1510 ACGGTGCTTGTGATGAGCGAGTGGACACGCGTCACTGATGATGATGATGATGATGAT 1569
 DB 964 ACGGTGCTTGTGATGAGCGAGTGGACACGCGTCACTGATGATGATGATGATGATGAT 1023
 QY 1570 TGCGCAGCTGCGCGCTGCGCGCTCAAGCAGCAGCGCGCGCGCGCGCGCGCGCGCG 1629
 DB 1024 TGCGCAGCTGCGCGCTGCGCGCTCAAGCAGCAGCGCGCGCGCGCGCGCGCGCGCG 1083
 QY 1630 CCGGCCATCAAGCAGCTGATTAAGATCTACAGCGCATGCTTACGCG 1675
 DB 1084 CCGGCCATCAAGCAGCTGATTAAGATCTACAGCGCATGCTTACGCG 1129

RESULT 4
 ID ABX75860 standard; cDNA, 955 BP.
 AC ABX75860;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DE Human Neu2 cDNA #3.
 XX
 XX Human; mouse; rat; gene; 89; neutralised polypeptide; Neu; NHR; cancer;
 KM neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;
 KM neurogenic function; transcriptional regulator; cell signaling; pain;
 KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signaling; axonal path-finding;
 KM neurodegenerative disease; cytoskeletal; antidepressant; analgesic;
 KM transducer; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 XX
 XX Homo sapiens.
 OS
 XX
 PN US2002132293-A1.

XX 19-SEP-2002.
 PD 14-MAR-2001; 2001US-0808387.
 XX 14-MAR-2001; 2001US-0808387.
 PP 14-MAR-2001; 2001US-0808387.
 PR 14-MAR-2001; 2001US-0808387.
 XX (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 XX Palm K, Timmusk T;
 XX WPI: 2003-174081/17.
 DR P-PSDB; ABUS6253.
 DR Novel neutralized polypeptide, Neu useful for controlling cell
 PT proliferation and calcium signaling induced transcriptional processes
 PT and treating depression, pain, anxiety, cancer and neurodegenerative
 PT diseases -
 XX
 PS Claim 6; Page 43; 75pp; English.
 XX
 XX The invention relates to a purified neutralised (Neu) polypeptide
 CC comprising at least one neutralised homology repeat (NHR) domain and a
 CC C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
 CC polypeptide. The polynucleotide is useful for constructing a transformed
 CC host cell that expresses a Neu protein. Neu has neurogenic function and
 CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signaling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signaling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signaling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.
 XX
 SQ Sequence 955 BP; 148 A; 355 C; 304 G; 148 T; 0 other;
 Query Match 34.5%; Score 577.2; DB 25; Length 955;
 Best Local Similarity 99.5%; Pred. No. 3.1e-83;
 Matches 579; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCAACAGCTGACACCGGACCTTGCACAGCCGAGCCCAACCGGCGCGCTCTTGGCC 60
 DB 1 ATGGGCAACAGCTGACACCGGACCTTGCACAGCCGAGCCCAACCGGCGCGCTCTTGGCC 60
 QY 61 ACCGGCGGTGCTGCG 120
 DB 61 ACCGGCGGTGCTGCG 120
 QY 121 TTCCACGCGCAGGCGCAAGAGCAGTGGCGCTGAGCAGCAGCTCGCGCGCGCGCA 180
 DB 121 TTCCACGCGCAGGCGCAAGAGCAGTGGCGCTGAGCAGCAGCTCGCGCGCGCGCA 180
 QY 181 CCGGCGCAACAGCTTCTGCAATGAGCTCACCTTTCACAGCGCGCCCATTCGGCTGTACAG 240
 DB 181 CCGGCGCAACAGCTTCTGCAATGAGCTCACCTTTCACAGCGCGCCCATTCGGCTGTACAG 240
 QY 241 CAGGTGCGGCTGCGCGCTGAGTGGCGCGCGCTGAGCGAGCGCGCGCTGCGCGCGCG 300
 DB 241 CAGGTGCGGCTGCGCGCTGAGTGGCGCGCGCTGAGCGAGCGCGCGCTGCGCGCGCG 300
 QY 301 TTCCACGCGCAGATTCGCTGCTCATGAGCGCGCAAGCATCCCAAGTACGCTGCGCG 360
 DB 301 TTCCACGCGCAGATTCGCTGCTCATGAGCGCGCAAGCATCCCAAGTACGCTGCGCG 360

QY 361 GACCTGTACACGCGCGCGGCTACTGTGGCCAAAGCACTGCCGAGAACTGGCGCTGGCGC 420
 DB 361 GACCTGTACACGCGCGCGGCTACTGTGGCCAAAGCACTGCCGAGAACTGGCGCTGGCGC 420
 QY 421 GACACGCTGTCTGCTCTACTGTGGCCGACCGCCGCGCTGTCTTCAACAGCTGAACGAC 480
 DB 421 GACACGCTGTCTGCTCTACTGTGGCCGACCGCCGCGCTGTCTTCAACAGCTGAACGAC 480
 QY 481 GCGGAGCGCGCTCTTCCACGCGCGCGGCTGGCGCGGCGCGCGCTGGCGCGCTCAT 540
 DB 481 GCGGAGCGCGCTCTTCCACGCGCGCGGCTGGCGCGGCGCGCGCTGGCGCGCTCAT 540
 QY 541 GATGTCTACGCGCATCACCGACGAGGTGACGCTTCTGAGAGC 582
 DB 541 GATGTCTACGCGCATCACCGACGAGGTGACGCTTCTGAGAGC 582
 RESULT 5
 ABQ43992/C
 ID ABQ43992 standard; DNA, 1016 BP.
 AC ABQ43992;
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30583.
 XX
 KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 KM SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 PD 07-MAR-2002.
 PF 01-SEP-2001; 2001WO-EP10074.
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIC-) EPIDEMICS AG.
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (1) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNPs) and (1) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ43992 represents genomic DNA sequences used to illustrate the

CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 1016 BP; 132 A; 154 C; 421 G; 306 T; 3 other;
 Query Match 25.8%; Score 432.4; DB 24; Length 1016;
 Beat Local Similarity 78.0%; Pred. No. 36-60;
 Matches 581; Conservative 0; Mismatches 159; Indels 5; Gaps 5;
 QY 578 AGAGCGCTTGTGCTGACACGCTGACGCGCGCGCTTCAAGCCGCTTCAAGCGCT 637
 DB 912 AAAAGCGCTTGTGCTGACACGCTGACGCGCGCGCTTCAAGCCGCTTCAAGCGCT 854
 QY 638 GCTTGTGCG 697
 DB 853 ACCTACCG 794
 QY 698 AGGTGTGTGCG 757
 DB 793 AAATTAATTAACCAAACTAAACCACTTG-ACTAAACCGCGCGCGCGCGCGCGCGCG 736
 QY 758 AGCG 816
 DB 735 AGCG 676
 QY 817 CGGCGCTACTGAGAGCG 876
 DB 675 CGGAGCTACTGAGAGCG 616
 QY 877 TCGGCG 936
 DB 615 TCGACCGAGCG 556
 QY 937 TCGG-AGCG 995
 DB 555 TCGGAAAGCG 496
 QY 996 GCG 1055
 DB 495 AACGAGCG 436
 QY 1056 CAAGAGCG 1115
 DB 435 CAACGAGCTACTGAGAGCG 376
 QY 1116 GCG 1175
 DB 375 GCG 316
 QY 1176 GCAAGCTGTCTGTGAGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
 DB 315 GCAAGCTGTCTGTGAGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 256
 QY 1236 GCAAGCTGTCTGTGAGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
 DB 255 GCAAGCTGTCTGTGAGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 196
 QY 1296 CGGTACCGTGTGAGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
 DB 195 CGGTAAATCCCGAGCG 171
 RESULT 6
 ABQ43993
 ID ABQ43993 standard; DNA, 1016 BP.
 AC ABQ43993;
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30584.
 KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 KM SNP; cell differentiation; db.
 XX Homo sapiens.
 OS
 PN WC0200218632-A2.
 PD 07-MAR-2002.
 PF 01-SEP-2001; 2001WO-BP10074.
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX (EPiG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 DR
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc.; particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's) and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ1310-ABQ4121 represent genetic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 1016 BP; 306 A; 421 C; 154 G; 132 T; 3 other;
 Query Match 25.8%; Score 432.4; DB 24; Length 1016;
 Best Local Similarity 78.0%; Pred. No. 3e-60;
 Matches 581; Conservative 0; Mismatches 159; Indels 5; Gaps 5;

QY 877 TCGCGGACCGCAAGTGGCTGCGACACGCGGCGCGACGCGCGACGCTGCTTC 936
 DB 402 TCGACGACCGCAAAATTAACCTACGACMAAGACCGAACGACGACGCTAATCTTC 461
 QY 937 TCCG-AGCGCGCGCTGCGCGCGCGCGAGACCTCTTCGTGAGAGTGGCCGTCCGCGCT 995
 DB 462 TCCGAANCCCGCTACGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
 QY 996 GCG 1055
 DB 522 AACGACG 581
 QY 1056 CAACGAGCTGCG 1115
 DB 582 CAACGAGCTGCG 641
 QY 1116 GCG 1175
 DB 642 GCG 701
 QY 1176 CGACGTGCTCTTGGGATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
 DB 702 CGACGTGCTCTTAAACATCAACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 761
 QY 1236 GCG 1295
 DB 762 GCG 821
 QY 1296 CGGTACCTGCTGAGTCCAGCGCTGCG 1320
 DB 822 CGATTAATCCCGACCGCGCGCTAG 846
 RESULT 7
 ID ABX75849 standard; cDNA; 1674 BP.
 XX
 AC ABX75849;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DB Human Neu1 cDNA #2.
 XX
 KM Human; mouse; rat; gene; 88; neutralised polypeptide; Neu; NHR; cancer;
 KM neutralised homology repeat; C3Hc4 RING-zinc finger domain; myogenesis;
 KM neurogenic function; transcriptional regulator; cell signalling; pain;
 KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signalling; axonal path-finding;
 KM neurodegenerative disease; cytoskeletal; antidepressant; analgesic;
 KM tranquiliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 XX
 OS Homo sapiens.
 XX
 PN US2002132293-A1.
 XX
 PD 19-SEP-2002.
 XX
 PF 14-MAR-2001; 2001US-0808387.
 XX
 PR 14-MAR-2001; 2001US-0808387.
 PA (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 XX
 PT Palm K, Timmusk T;
 XX
 DR WPI; 2003-174081/17.
 DR P-PSDB; ABUS6242.
 XX
 PT Novel neutralized polypeptide, Neu useful for controlling cell

PT proliferation and calcium signaling induced transcriptional processes
PT and treating depression, pain, anxiety, cancer and neurodegenerative
PT diseases -

PS Claim 6; Page 21-22; 75pp; English.

XX The invention relates to a purified neuralised (Neu) polypeptide
XX comprising at least one neuralised homology repeat (NHR) domain and a
CC C3H4 RING-zinc finger domain, and a polynucleotide encoding the
CC polypeptide. The polynucleotide is useful for constructing a transformed
CC host cell that expresses a Neu protein. Neu has neurogenic function and
CC functions as a transcriptional regulator, as a calcium-signal transducer,
CC in cell signalling and synaptogenesis, in memory learning,
CC tumorigenesis, myogenesis and development of other organ systems and is
CC related to repair and regeneration after injury to the central nervous
CC system. Neu interacts with neurogenic genes, proteins implicated in
CC nuclear transport and with Parkin-like proteins. Neu has the potential to
CC interfere with inductive tissue interactions, cytokine signalling, RNA
CC processing, early immediate responses, death of specific cell
CC populations, nuclear hormone receptor signalling and axonal path-finding.
CC The Neu family of proteins presents a set of diagnostic and therapeutic
CC targets, especially for treating cancer. Manipulating Neu expression and
CC function is useful in controlling a variety of diseases including
CC depression, pain, anxiety, neurodegenerative diseases and cancer.
CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
CC invention.

XX Sequence 1674 BP; 257 A; 634 C; 510 G; 273 T; 0 other;

Query Match 24.4%; Score 409.2; DB 25; Length 1674;
Best Local Similarity 56.6%; Pred. No. 1.4e-56;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY 99 GGTCTCGGCGAGGCGCCCGCTTCCACCGCGAGCGCAAGCAAGAAAGTGGGCTGA 158
DB 120 GCTCCAGACACCCCTGCTCTTCCACCGCACCAAGGCTCCCAATCTCATGGA 179
QY 159 CGGCACTCGCGCGCGGACCAACAGCGGCAACAGCTTCTGCAATGCGTCACTTCAACA 218
DB 180 CTTCAAGCCACAAAGGCTGTCAAGAGGAGCGCAAGCTTCTGCAAGCCATCTTACGAA 239
QY 219 GCGGCCCATCCGCGCTGTACAGAGAGTGGCGCTGCGCTGAGGCGCGTGGCGCTG 278
DB 240 CGCGCGGCTCTCATCTACAGAGCAAGTCAAGCTGAAGATCACCAAGAGAGTGTG 299
QY 279 GAGCGCGCGCTGCGCTTCCACCGCGCAACATCGGTGCTCATGAGCGCCAGGA 338
DB 300 GAGCGGCGCGCTGCGCTTCCACCGCAAGAGACCGGTCCCGCATCCACCTGATCTC 359
QY 339 CATCCCAAGTACGCGTCCGAGCGGAGCTGAGCGCGGCGGCTACGTGAGCCAAAGGACT 398
DB 360 GCTGCCCAAGTACGCTTGGCCGACCTGGTGTCCAGAGTGGCTTTTGGGCGCAAGCGCT 419
QY 399 GCCCGAAGAACTGCGCTGCGGACACGCTGCTGAGCTACTGAGCCGACCGCGCG 458
DB 420 GCTGAGAGGTTTGCAATGAGGGAACAATCATCGATTCTGGGTGACAAAGAGGCG 479
QY 459 CGTGTCTACAGGTGAGGAGCGAGCGGCTCTTCACTGCGGCGTGGCGCTG 518
DB 480 TGTCTTCAACCGGATCAAGCACTCGGCTGTATGCTGTTCTTAAAGGGGTCCGACGGC 539
QY 519 CGGCGCGCTTGGGCGCTCATGATGTCTACGGCATCACCGAGAGGTGAGCTTCTGGA 578
DB 540 CGACCGGCTCTGGGCGCTGGTGAAGTCTACGGCTCTACCGGGGGGCTCCAGTGTGA 599
QY 579 GAGCGGCTTCTGAGACGCTGACGCGCGCGCTTCAAGCAAGCCCGCTTACGGCGCT 638
DB 600 TAGCGAGCTGG-----TGTCCGGAAGTGTGCGGCGCGCTTCTTCAAGCGCGCT 650
QY 639 CTTGCGCGCGGAGCAAGCGCGGCACTTTCAGCAACAAAGAGCTGAGAAACAACA 698
DB 651 GCGGCGGCGGTGCTGCGGCGCGAGGCGGAGCAAGCGCGCTCTCGGTG----- 699

QY 699 GGTGAGGCAAGCTGAGGCACTGAGGCGCTGAGGCGCGCGCGCGCGCAAGCGCA 758
DB 700 -----AGCTATGAGCACTCAAGTGGCGGCGCGAGCGGAGCGCGCG 749
QY 759 GCGCGCGCGCGCGCATTTCCGTGAGGAGCGCGCGCGCGCGCGCGCTGTCGCC 818
DB 750 GCG 809
QY 819 GCGCGTATGAGAGCGGAGCTGCGCTTCCAGCAACAGCGGAGCGCG---AGTGAAGCT 875
DB 810 GAGCGAGCTTCAAGCGGAGCTGCGCTTCCAGCGCGCGCGCGCGCGCGCGCGCG 869
QY 876 GTGCGCGAGCGCAAGTGTGCTGCGGACCGCGGCGCGAGCGGCGCGCGCGCTGCT 935
DB 870 CTTCAAGAGAGAGAGCGGTGCGCGCGGTGAGAGCAAGGCGCGCGCGCGCGCTGCT 929
QY 936 CTTCAAGAGAGCTG 995
DB 930 CACCAAGCG 989
QY 996 GCG 1055
DB 990 GCG 1049
QY 1056 CAAGAGCTGCG 1115
DB 1050 GAGCGAGCTGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
QY 1116 GCG 1175
DB 1110 CGCGGTGCG 1169
QY 1176 CGAGGTGCTCTGAGGCACTCAAGCGGCGGTGCGCGCGCGCGCGCGCGCGCG 1235
DB 1170 CGAGGTGCACTCAAGCGCAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
QY 1236 GAGGAGCTGTGAGGCTTCTTGTGCGCGGTGCGCGCGCGCGCGCGCGCGCG 1295
DB 1230 GAGGCGCGCTTGTGAGTGTCTTGTGCGCGCGCGCGCGCGCGCGCGCGCG 1289
QY 1296 CGGTACCTGTGAGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1346
DB 1290 CTTCACTATCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1349
QY 1347 -----CGGCTCCAGAGAGATGATGATTTAGATTAAGCTTCAAGTCAACAGTC 1397
DB 1350 AACCTGCGCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1409
QY 1398 CTTCTGCGGATCTGAGTCAATCCCTGTGAGCGGCGCGCGCGCGCGCGCGCG 1457
DB 1410 TGGCGCTGTGAGTGTGCTGTGAGTGTGAGCAAGCGCGCGCGCGCGCGCGCG 1469
QY 1458 GTTCCCGCTGTGTCTCCCAACCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1517
DB 1470 GTTCCCGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1520
QY 1518 CTTCAAGTGAAGTGTGAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1577
DB 1521 CTAAGAGAGCGGAGTGTGAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1580
QY 1578 CTTGCGCGCTGCGGCTCAAGAGCAAGGCGCGCGCGCGCGCGCGCGCGCGCG 1637
DB 1581 CTTGCGCGCTGCGGCTCAAGAGAGCTGTGAGCGCGCGCGCGCGCGCGCGCG 1640
QY 1638 CAAGAGCTCATTAAGTCTACAG 1661
DB 1641 CAAGAGCATCATTAAGACTTACCG 1664

RESULT 8
ABX75848
ID ABX75848 standard; cDNA; 1725 BP.
XX

AC ABX75848;
XX 31-MAR-2003 (first entry)
XX Human Neu1 cDNA #1.
DE
XX
KM Human; mouse; rat; gene; ss; neutralised polypeptide; Neu; NHR; cancer;
KM neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;
KM neurogenic function; transcriptional regulator; cell signalling; pain;
KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
KM tumorigenesis; organ development; central nervous system; depression;
KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
KM inductive tissue interaction; cytokine signalling; axonal path-finding;
KM neurodegenerative disease; cytostatic; antidepressant; analgesic;
KM tranqulliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
XX
XX Homo sapiens.
OS
XX US2002132293-A1.
PN
XX 19-SEP-2002.
PD
XX 14-MAR-2001; 2001US-0808387.
PF
XX 14-MAR-2001; 2001US-0808387.
PR
XX 14-MAR-2001; 2001US-0808387.
PX
XX (PALM/) PALM K.
PA (TIMM/) TIMMUSK T.
PI Palm K, Timmusk T;
XX
XX WPI; 2003-174081/17.
DR P-PSDB; AB056241.
PT
XX Novel neutralized polypeptide, Neu useful for controlling cell
PT proliferation and calcium signaling induced transcriptional processes
PT and treating depression, pain, anxiety, cancer and neurodegenerative
PT diseases -
XX
XX Claim 6; Page 19-20; 75pp; English.
PS
XX The invention relates to a purified neutralised (Neu) polypeptide
XX comprising at least one neutralised homology repeat (NHR) domain and a
XX C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
XX polypeptide. The polynucleotide is useful for constructing a transformed
XX host cell that expresses a Neu protein. Neu has neurogenic function and
XX functions as a transcriptional regulator, as a calcium-signal transducer,
XX in cell signalling and synaptogenesis, in memory learning,
XX tumorigenesis, myogenesis and development of other organ systems and is
XX related to repair and regeneration after injury to the central nervous
XX system. Neu interacts with neurogenic genes, proteins implicated in
XX nuclear transport and with Parkin-like proteins. Neu has the potential to
XX interfere with inductive tissue interactions, cytokine signalling, RNA
XX processing, early immediate responses, death of specific cell
XX populations, nuclear hormone receptor signalling and axonal path-finding.
XX The Neu family of proteins presents a set of diagnostic and therapeutic
XX targets, especially for treating cancer. Manipulating Neu expression and
XX function is useful in controlling a variety of diseases including
XX depression, pain, anxiety, neurodegenerative diseases and cancer.
XX Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
XX invention.
XX
SQ Sequence 1725 BP; 268 A; 658 C; 520 G; 279 T; 0 other;

Query Match 24.4%; Score 409.2; DB 25; Length 1725;
Best Local Similarity 56.6%; Pred. No. 1.4e-56;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY 99 GGTCTGGAGGAGGCGCGCTTCACGCGGAGGCGCAAGAAAGTCGCGCTGGA 158
DB 171 GCTCCACGACGCGCGCTGCTTCACCGCGACCAAGGCGCTCCGAGATCTCATGGA 230
QY 159 GCGCACTGCGCGCGCGGCAACGCGGCAAGCACTTTCGCAATGCGTCACGTTACGGA 218

DB 231 CTTAGGCAAGAGGCTGTCAAGAGGAGGAGGCGGAGCTTTCGAGAACCATTCACGAA 290
QY 219 GCGGCAATCCGCTGTACAGAGAGTGGCGGCTGGTGGCCCTGGCTG 278
DB 291 CCGCCCGGCTCTCATCTACAGAGCAAGTCAAGCTCAAGATCAACAAAGAGTGTCTG 350
QY 279 GAGGAGGCGGCTTGGCTTCCGCTTCAACCGGAGAGTCCGTGCTCAAGAGGCGGAG 338
DB 351 GAGGAGGCGGCTTGGCTTCCGCTTCAACCGGAGAGTCCGTGCTCAAGAGGCGGAG 410
QY 339 CATCCCAAGTACGCTTCCGCTTCAACCGGAGAGTCCGTGCTCAAGAGGCGGAG 398
DB 411 GCTCCCAAGTACGCTTCCGCTTCAACCGGAGAGTCCGTGCTCAAGAGGCGGAG 470
QY 399 GCGGAGAGCTTGGCTTCCGCTTCAACCGGAGAGTCCGTGCTCAAGAGGCGGAG 458
DB 471 GCTTGAAGAGTTCCTCAATGAGGAGCAATCATGCTTCTGGGTGAGCAAGAGGCGG 530
QY 459 GGTGTTCTACAGCGTGAACGAGCGGAGCGGCTTCTCACTGGCGGCTGGCGGCTG 518
DB 531 TGTCTTCAAGCAATCAAGCACTCGCTGTTATGCTTTCAGCGGAGGCTCGACGAGC 590
QY 519 GCGGCGGCTTGGGCGCTCATTTGATGTTCAAGGATCAAGAGAGTGCAGCTTGTGA 578
DB 591 GAGCCGCTTGGGCGCTTGGGAGCTTCAAGGCTTCAAGGAGGCTTCAAGGCTTGA 650
QY 579 GAGGCTTGGCTTGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTGA 638
DB 651 TAGGAGCTGG-----TGTCCCGGAGCTGTGGGCGGCGCTTCAAGGCTT 701
QY 639 CTTGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
DB 702 GCGGCGGCTTGGCTTGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTGA 750
QY 699 GATGATGAGCAAGTGGGAGCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
DB 751 -----AGCTTATGAGCACTTCAAGTGTCCGCGGCGGAGGAGGAGGAGGAGGAG 800
QY 759 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
DB 801 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 860
QY 819 GCGGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
DB 861 GCGGAGGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 920
QY 876 GTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935
DB 921 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 980
QY 936 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 995
DB 981 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1040
QY 996 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1055
DB 1041 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1100
QY 1056 CAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1115
DB 1101 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1160
QY 1116 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1175
DB 1161 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220
QY 1176 GAG 1235
DB 1221 GAG 1280
QY 1236 GAG 1295

Db 1281 GCAGCGCTTGGATGCTCTTGGCTGACGCGGACCATCACGAGATCCGATCCTCGG 1340
 1296 CGGTACCTGAGTCCAGGCGCTTGCACGACATCCATCAGGGTCCCTGAG----- 1346
 Db 1341 CTCGATATCTGGCGGAGCGGAGTATCCGATCATCCCTGCTCCCTGCTCCAGCGC 1400
 1347 -----CGGCTCCAGAGCATAGTATTCAGATGATGACCTTCAGTCAACGAGTC 1397
 1401 AACCTGCGCCAGTGGCCCTGGGAGCGCGCTGTGACCCCTTCTCAGGACGTCAGCTC 1460
 Qy 1398 CTCCTGGGATCTGATCATCCTGGTGAAGGCGCCCAAGTCCCGCTGAGCCCCCGGT 1457
 1461 TGGCCCTTGGGAGTCTGTGCTGGTGAACAGCCCCCAATTGCGAGTGAAGCTGCGCA 1520
 Qy 1458 GTCCCGGCTTCTCCCAACCGGAGCGGAGGATCAAGAAATGGGAGTGAAGGAGTGA 1517
 1521 GTCCGAGTGAACCCAGGCTTGGGCAAGT-----GAGGATGAGTGAACCATTTG 1571
 Qy 1518 CTTGATGGGAGTGAACAGGTCATTAACAGTGTGACACATGTGCTGTGCAAG 1577
 1572 CATGAACACGGGCTGGAACGGTCAATCTACATGTGGCCATGTGCTGTACCG 1631
 Db 1578 CTGCGGCTGCGGCTCAAGGACAGGCGCGGCTGTGCTGCCATCTGCGAGGCGCAT 1637
 1632 CTGTGGCTGCGGCTCAAGAGGCTGTGACAGCTGTGCTGCCATCTGCGCGCGCCAT 1691
 Qy 1638 CAAGGAGTCATTAAGTCTACAG 1661
 1692 CAGGAGATCATCAAGACTACCG 1715
 Db

RESULT 9.

AAX77135
 AAX77135 standard; DNA; 2207 BP.

AAX77135;

03-AUG-1999 (first entry)

Human nerve mutation factor protein encoding DNA.

Nerve mutation factor; chromosome 10; glioma; tumour suppressor;
 brain tumour; astrocytoma; gene therapy; human; 88.

Homo sapiens.

MO925827-A1.

27-MAY-1999.

24-AUG-1998; 98MO-JP03737.

14-NOV-1997; 97JP-0313211.

(SUME) SUMITOMO ELECTRIC IND CO.

Nakamura H, Nakata M, Saya H, Yoshida M;

WPI; 1999-347474/29.

P-PSDB; AAY21558.

Human gene on chromosome 10 homologous to Drosophila neuralized
 gene, useful in the diagnosis and gene therapy of brain tumors

Claim 9; Page 49-56; 78pp; Japanese.

The invention relates to a protein which is a nerve mutation factor and
 is the expression product of a gene located on chromosome 10. The gene
 is in a region frequently deleted in highly malignant gliomas. Sequences
 (AAX77135 and AAX77136) encoding human and mouse nerve mutation factors
 (AAY21558 and AAY21559) are provided. The protein is believed to have
 tumour suppressor activity. Polynucleotide sequences and antibodies to
 the protein are diagnostic reagents for highly malignant brain tumors

CC such as astrocytoma where chromosome 10 deletion commonly occurs. The
 CC gene may also be used for gene therapy of such tumors. The present
 CC sequence represents a DNA encoding the human nerve mutation factor.

SQ Sequence 2207 BP; 340 A; 872 C; 673 G; 322 T; 0 other;

Query Match 24.4%; Score 409.2; DB 20; Length 2207;
 Best Local Similarity 56.6%; Pred. No. 1.4e-56;
 Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

Qy 99 GGTCTTGGGAGGAGGCGCGCTTTCACGCGGACCAAGGCAAGAGCTGGAG 158
 581 GTCTCCAGACGCGGCTGTCTTCACCGGACCAAGAGGCTCCCAATCTCATGGA 640
 Qy 159 CGGCACTGCGCGCGGACCAAGGCGGCAAGGCTTTCAGATGAGGCTTTCAGCA 218
 641 CTTAGGCAAGAGGCTGTCAAGAGGAGGCGGAGCTTTCAGGATCACTTTCAGCA 700
 Db 219 GCGGCCCATCCGACTGTACAGACAGTGGGCTGCGCTGTGGCCCTGAGCTG 278
 701 CGGCGGCTCATCTACAGCAAGTCAAGGCTGAAGATCAAGAAAGAGTGTGCTG 760
 Qy 279 GAGGCGGCGCTGTGCGCTTTCAGCGGAGACATCCGTGTCTATGAGGCCAGGA 338
 761 GAGCGGCGCTGTGCGGCTTTCAGCGGAGACATCCGTGTCTATGAGGCCAGGA 820
 Qy 339 CATCCCAAGTACGCGCTGTGCGGAGCTGTGCTGAGCGGCGGCTTTCAGGCAAGCACT 398
 821 GCTGCCAAGTACGCGCTGTGCGGAGCTGTGCTGAGCGGCGGCTTTCAGGCAAGCACT 880
 Db 399 GCGCGGAGACTGTGCGGCTGTGCGGAGACATCCGTGTCTATGAGGCCAGGA 458
 881 GCTGAGGAGTGTGCAATGAGGAGCAATCATGCAATTCGAGGTGGAAGAAAGGCGG 940
 Qy 459 CGTGTCTACAGCTGAAGAGGAGGCGGCTTTCAGCGGAGCTTTCAGGCAAGCACT 518
 941 TGTCTTTCACAGCTGAAGAGGAGGCGGCTTTCAGCGGAGCTTTCAGGCAAGCACT 1000
 Qy 519 CGGCGGCTGTGCGGCTGTGCGGAGCTGTGCTGAGCGGCGGCTTTCAGGCAAGCACT 578
 1001 CGAGCGGCTGTGCGGCTGTGCGGAGCTGTGCTGAGCGGCGGCTTTCAGGCAAGCACT 1060
 Qy 579 GAGCGGCTTTCAGCTGAAGAGGAGGCGGCTTTCAGCGGAGCTTTCAGGCAAGCACT 638
 1061 TAGCGAGCTGTG-----TGTCTCCGAGTGTCTGCGGCGGCTTTCAGCGGCGCT 1111
 Db 639 CTTGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
 1112 GCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1160
 Qy 699 GGTGTGCGCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
 1161 -----AGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1210
 Qy 759 CGCGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
 1211 GCGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
 Db 819 GCGCGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
 1271 GCGCGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1330
 Qy 876 GTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935
 1331 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
 Qy 936 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 995
 1391 CACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450
 Qy 996 GCGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1055
 1451 CGCGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510

QY 819 GAGCGTACTGAGAGCGGACCTGCGCTTCCACGCAACACGCGGCGCCG--ACGTGAGCCT 875
 DB 777 GGGCGAGCTGACGCGGACCTGGCGCTTCCACGCGCCTTCCGCGCGCGCGACGCTCCGCGAT 836
 QY 876 GTGGCGGACCGGCAAGAGTGGCGCTTGGCGACCGCGCGCGCGCGCGCGCGCGCGCTT 935
 DB 837 CTTGAGACGACGAGCGGTGGCGCGCTTGGAGACCGGCGCGCGCGCGCGCGCTT 896
 QY 936 CTCGAGCT 995
 DB 897 CACGAGCG 956
 QY 996 GCG 1055
 DB 957 GGGCGGAGCG 1016
 QY 1056 CAACGAGCTGCG 1115
 DB 1017 CCGCGAGCTGCG 1076
 QY 1116 GCG 1175
 DB 1077 TCCG 1136
 QY 1176 CGAGCTGCTCTGGGCGATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
 DB 1137 AGAGCTGACCTGAGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1196
 QY 1236 GCGAGCGCTGCG 1295
 DB 1197 GCGAGCGCTGCG 1256
 QY 1296 CGGTAACCTGCGAGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1348
 DB 1257 CTCGACCATATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
 QY 1349 -----GCTCCAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1397
 DB 1317 AACCTGACCGAGTGCCTGGGCGATCCGCTCTGAGACCGCGCTGCGAGCACTGCGGTT 1376
 QY 1398 CTCTCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
 DB 1377 TGGGCG 1436
 QY 1458 GTCCCG 1517
 DB 1437 GCCACCGGAGTACCCAGAGTCTGGGCGAG-----TGAGATGATGATGATGATGAT 1487
 QY 1518 CTTGATGAGCGAGTGAACGCGTCACTGATGATGATGATGATGATGATGATGATGATGAT 1577
 DB 1488 CTATGAAACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1547
 QY 1578 CTGCG 1637
 DB 1548 CTGCG 1607
 QY 1638 CAAGGACGTGATTAAGATCTACAG 1661
 DB 1608 CAAGGACATCATCAAGACTTACCG 1631

RESULT 11
 ABX75852
 ID ABX75852 standard; cDNA; 1698 BP.
 AC ABX75852;
 DT 31-MAR-2003 (first entry)
 XX Mouse Neu1 cDNA #2.
 DE Human; mouse; rat; gene; 88; neutralised polypeptide; Neu; NHR; cancer;
 KW neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;

KM neurogenic function; transcriptional regulator; cell signalling; pain;
 KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signalling; axonal path-finding;
 KM neurodegenerative diseases; cytotoxic; antidepressant; analgesic;
 KM tranquiliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 OS Mus sp.
 XX US2002132293-A1.
 PN 19-SEP-2002.
 XX 14-MAR-2001; 2001US-0808387.
 PP 14-MAR-2001; 2001US-0808387.
 PR 14-MAR-2001; 2001US-0808387.
 PA (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 PI Palm K, Timmusk T,
 XX DR WPI, 2003-174081/17.
 DR P-PSDB; AB056245.
 XX PT Novel neutralized polypeptide, Neu useful for controlling cell
 PT proliferation and calcium signaling induced transcriptional processes
 PT and treating depression, pain, anxiety, cancer and neurodegenerative
 PT diseases -
 XX PS Claim 6; Page 28-29; 75pp; English.
 CC The invention relates to a purified neutralised (Neu) polypeptide
 CC comprising at least one neutralised homology repeat (NHR) domain and a
 CC C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
 CC polypeptide. The polynucleotide is useful for constructing a transformed
 CC host cell that expresses a Neu protein. Neu has neurogenic function and
 CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75846-ABX75864 represent Neu polynucleotides of the
 CC invention.
 CC
 XX Sequence 1698 BP; 277 A; 608 C; 514 G; 299 T; 0 other;
 SQ
 Query Match 23.4%; Score 391.6; DB 25; Length 1698;
 Best Local Similarity 55.9%; Pred. No. 8.7e-54;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

QY 99 GATCTGGGAGGCGCGCGCTTCCACGCGCGCAAGGCAAGCAAGCAAGCTGCGCTGCA 158
 DB 120 GCTCCCGGCGACCGCGCTGCTTCCACCGCGCACTAAGGAGGCTCCCAAGTCTCAATGGA 179
 QY 159 CGGCGACTGCGCGCGCGCGCAACGCGCGCAACGCTTCTGCAATGAGGTCAAGTTACGCA 218
 DB 180 CTTGAGCACAGGCGCGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 239
 QY 219 GCGGCGCATCCGCGTGAACGAGAGTGGCGCTGCGCGCTGAGTGGCGCGCGCGCGCGCTG 278
 DB 240 CGGCGCGGCTCATCTAAGAGAGTCAAGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 299
 QY 279 GAGCGGCGCGCTGCGCTGCGCTTCCACGCGCGCAAGTCCGCTCATGAGCGCCAGGA 338

Db 300 GAGCGGGGCCCTGCGCACTTGGCTTCAACGAGCAAGAACCTTCCCGATCCACCCCGACTC 359

QY 339 CATCCCAAGTACGCTTCGCGCGAAGCTGGTCAACGCGGCGCGGCTACTGGGACCAAGGCACT 398

Db 360 GCTGCCCAAGTACGCTTCGCGCGTGAACCTGGTGTCTCAGATGGGCTTCTGGGCCCAAGCAATT 419

QY 399 GCCCGAGAACCTGGGCGTGGCGGAGCAACGGTGTCTGGCTTCTGGGCGCGACCGCACCGGCGG 458

Db 420 GCTGAGGAGTTTGGCAACAGAGGCAACATCATTTGCTTCTGGTGGAGAACAGAGGAGCGG 479

QY 459 CGTGTCTTACAGCGTGAACGAGCGGAGCGGAGTCTTCACTGGGAGGCTGGCGCTGGG 518

Db 480 CGTCTTCTACCGGATCAATGAGTCAAGCTCTATGCTTTCTTTCAGTGGGAGTCCGAGCGGT 539

QY 519 CGGCGCGCTCTGGGCGCTCATTTGATGTCTACGCGCATCAACGAGAGGTCAAGTTCTTGA 578

Db 540 GGAACCGGCTCTGGGCGCTGGTGAAGTCAAGCGGCTCAACGCGGGGAGTCAAGCTGTAGA 599

QY 579 GAGCGCCTTTCGCTGACAGCTGACGCGCGCGCGGCTCAAGCGGCGGCTTACGCGCTG 638

Db 600 CAGCAGAGCTGG-----TGCTGCCCACTGCTCGCGCGCGCTCTCTTCAACCGCGCT 650

QY 639 CCTGCGCGCCAGCAGCGACGAGCGCGGCGCAACTTCCAGCAACAGAGCTGAGAACAGCA 698

Db 651 GCGGCGGCGCTGCTGCGGTGCGAGGCG-----GATGAGCGCG 689

QY 699 GGTGTGGCCCAAGCTGAGGCACTTGGCGCTGGGCGCGGCGCGCGCGCCCAACCGCAACCGA 758

Db 690 CCTGTGCGTGAAGCTGTGGGACCTCAACGTTGCGGAGCGAGCGAGCGAGAGAGCGGCGCAC 749

QY 759 CGCGCGGCGCGCGCCCATTTCCGTGGGAGGCGCCGCTGAGCGCGCGCGGCGCGCTGTCGC 818

Db 750 GCCTGCCGAGCTGCCGATCCCGCAGAACCTCGCTCAATTCTCAGCAGCGCGCGCTGCGC 809

QY 819 GCGCGTACTGAGAGGCGGACCTGGGCTTCCAGCGCAACAGCGGCGCGG---AGTGAGCGCT 875

Db 810 GCGCGAGCTCAAGCGCGAGCTGCGCTTCCACGCGCTTTCGCGCGGCGCGCACGTCGCAAT 869

QY 876 GTGCGCCGACCGCAAAAGTGGCTTGGCGCACCGCGGCGCCGAGCGGCGGACGCTGGTCTT 935

Db 870 CCTGACCAAGAGCAGTGGGCGGCTCTGGAGACAGCGGCGCGAGCAGAGCGGCGCTGATCTT 929

QY 936 CTCCGAGCGCGCGCTGCGGCGCGGCGAGAGCTTTCTGTGAGTGGGCGCTTCCGAGGCT 995

Db 930 CACCAAGCGCGCGCTGTGAGCGGTGGCGAGCAATCTTCAATCAAGGTCAAGCGGCTCGGCGG 989

QY 996 GCGCGCGCGCGCGCGCTTGGGCAATCAAGTGTGGAGACCGCGGCGGTGTCAAGGCT 1055

Db 990 GGGGAGAGCGGCGCGGCTGTCTTCCGAGGTCAACAGTGTGAGCCCTGGGACGCTGGGCGC 1044

QY 1056 CAACGAGCTGCGCGCGGACCCGAGACGCGGTGCTCGACGCGCAAAAGATTACTGGGTGGGCG 1111

Db 1050 GCGGACCTTGCTTTCAGGCTCCGAGGCGCTTGTGAGCCGCAAGAGATTTCTGGCGGTGTG 110

QY 1116 GCGCGCGCGGCGCGCTTGGCGGAGCGGCGGCGAGCGGCTCAAGCTTCAACGCTGCGGCGCGG 117

Db 1110 TCGGTGGCGCGGCGCTTGGCAACAGGCGGAGCAATCTGGGCGCTGGGTGTCAAGCGGAGCG 116

QY 1176 CGACGTGCTCTTGGGATCAACAGGCGGTCCGCGCGGCGCGCTGTGTGGTGTGACAGCAC 123

Db 1170 AGAGCTGCACTGAGTCAACAGCGCGCGGCGCGGCGGCGAGCTGTGTGGATGCTCTC 122

QY 1236 GCGGCGGCTTGGGCGCTTTCAGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 129

Db 1230 GCGGCGGCTTGGGAGTCTTTCAGCTTGAATGGCGGCAATCAAGAGGTCCGATCTCTCG 128

QY 1296 CGGTACCTTGAAGTCAAGCGCTTGGCAACAGCTCCATCAAGGATCC----- 134

Db 1290 CTCACACATCAATGATGAAAGGAGGTGGCCCATCTCCCTGTGTACCTGTTCACATCC 134

QY 1343 -----TCAGCGGCTCCCAAGAGAGATGATTCAGATATGACTTCAAGTGTCAACCAATC 139

Db	1350	AACGTCAACCCAGTGCCTCGGGACATCCGGCTTGTCTGAACCCCTTGCTCAGACACTCGGGTTC	1409
Qy	1358	CTCTCTGGGCACTGTAGTATCATCCCTGGTGTGACGGCCCCCAGCTCCCGCTGAGACCCCCCGGT	1457
Db	1410	TGGGCCCCCTAGGTGACTCTGTGTGAGAGGACAGCCCCCAATTCACTGTGACCTGGCCGA	1469
Qy	1458	GTCCCGCCGTGTCTTCCCAACCGAGCCGGACGAGCATCAAGATGAGCGAGTGCACGGTGTG	1517
Db	1470	GTCAACCGGTGACCCCAAGTCTTGGGCCAG-----TGAATATGATATCAACCATTTG	1520
Qy	1518	CTTGATGTCGAGGTGTGACACAGGTCATCTACACGTGTGTGACACATGTGCTGTGGCAG	1577
Db	1521	CTAAGAACACCGAGGTGATACAGTATCTACACGATGTGGCCACATGTGCTGTGTATCTC	1580
Qy	1578	CTGGGCGCTGGCGCTCAAGCGACAGCGCCGGGCTGTGCTGCTCCCATCTGCGCGGCCCAT	1637
Db	1581	CTGTGTGCTCTGGCGCTCAAGAGGCGCTGTGACGCGCTGTGCGCCCATCTGCGGCCCAT	1640
Qy	1638	CAAGGAGCTCATTAAGATCTACAG	1661
Db	1641	CAAGGACATCATCAAGACTACCG	1664
RESULT 12			
ID	ABX75851	standard; cDNA; 1749 BP.	
XX	ABX75851;		
XX	AC		
XX	DT	31-MAR-2003 (first entry)	
XX	DE	Mouse Neu1 cDNA #1.	
XX	KM	Human; mouse; rat; gene; s9; neutralised polypeptide; Neu; NHR; cancer;	
KM	KM	neutralised homology repeat; CHC4 RING-zinc finger domain; myogenesis;	
KM	KM	neurogenic function; transcriptional regulator; cell signalling; pain;	
KM	KM	calcium-signal transducer; synaptogenesis; memory learning; anxiety;	
KM	KM	tumorigenesis; organ development; central nervous system; depression;	
KM	KM	neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;	
KM	KM	inductive tissue interaction; cytokine signalling; axonal path-finding;	
KM	KM	neurodegenerative disease; cytostatic; antidepressant; analgesic;	
KM	KM	tranquilliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.	
XX	OS	Mus sp.	
XX	PN	US200232293-A1.	
XX	PD	19-SEP-2002.	
XX	PP	14-MAR-2001; 2001US-0808387.	
XX	PR	14-MAR-2001; 2001US-0808387.	
XX	PA	(PALM/) PALM K.	
XX	PA	(TIMM/) TIMMUSK T.	
XX	PI	Palm K, Timmusk T;	
XX	PI		
DR	DR	WPI: 2003-174081/17.	
DR	DR	P-P8DB; ABUS6244.	
PT	PT	Novel neutralized polypeptide, Neu useful for controlling cell	
PT	PT	proliferation and calcium signaling induced transcriptional processes	
PT	PT	and treating depression, pain, anxiety, cancer and neurodegenerative	
PT	PT	diseases -	
PS	PS	Claim 6, Page 25-26; 75pp; English.	
XX	XX	The invention relates to a purified neutralised (Neu) polypeptide	
CC	CC	comprising at least one neutralised homology repeat (NHR) domain and a	
CC	CC	CHC4 RING-zinc finger domain, and a polynucleotide encoding the	
CC	CC	polypeptide. The polynucleotide is useful for constructing a transformed	
CC	CC	host cell that expresses a Neu protein. Neu has neurogenic function and	

CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.

SQ Sequence 1749 BP; 285 A; 630 C; 526 G; 308 T; 0 other;

Query Match 23.4%; Score 391.6; DB 25; Length 1749;

Best Local Similarity 55.9%; Pred. No. 8.6e-54;

Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

QY 99 GGTCTGGAGCGAGCGCGCTTCCACGCGCAAGGCCAAGCAAGAACTGCGGCTGGA 158
 DB 171 GCTCCGCGCACCGCTGCTTCCACCCCACTAAGGGCTCCAGATCTCTATGGA 230
 QY 159 CGGCACTCGCGCGCGCGCAACAGGCGCAACAGCTTCTGATATGGGTCTCAAGCA 218
 DB 231 CTTACAGCAAGAGCGCTCAAGAGGCAAGGCGCAAGCTTCTGCAATGCACTTCAGTAA 290
 QY 219 GCGGCCATCCGCGCTGACAGAGAGTGGCGCTGCGCTGAGCGCGTGGCGCTGCTG 278
 DB 291 CGCGCGGTGCTCATCTACAGAGAGTCAAGCTGGAATCAACAAAGAAATGCGTGG 350
 QY 279 GAGCGCGCGCTGCGCTTCCACGCGCAAGATCGCTCATGAGCGCCAGGA 338
 DB 351 GAGCGGGCGCTCGACTGCTGCTTCAACAGCAAGAGACCTTCCGCAATCCACCGATC 410
 QY 339 CATCCCAAGTAGGCGCTGCGCGCAAGCTGACGCGCGCGGGCTAATGCGGCAAGGACT 398
 DB 411 GTCGCCAAGTAGCGCTGCGCTGACGCTGAGTCTCAAGAGTGGCTTCTGCGGCAAGACT 470
 QY 399 GCCCGAAGACTGCGCTGCGCGCAACAGGTGCTGCTTACGCGCGCAAGCGCGCG 458
 DB 471 GCTGAGAGATTTGGCCAAAGAGGCAACATTCGCTTCTGAGTGAACAAGAGCGCG 530
 QY 459 CGTGTCTACAGCGTGAACAGCGCGAGCGGTGCTTCTCACTGCGCGCTGCGCTGG 518
 DB 531 CGTCTCTACCGGATCAATGATGATGCTGCTATGCTTTCTTCAATGGGGTCCGAGCG 590
 QY 519 CGCGCGCTGCGCGCGCTCATATGATCTAAGGATCAACGAGAGGTGAGCTTCTGGA 578
 DB 591 GAGACCGCTGCGCGCTGAGTGAAGTCTAAGCGCTTCAACGCGGGGTCTCAAGCTGAGA 650
 QY 579 GAGCGCGCTGCTGAGCAAGCTGAGCGCGCGCGCTGAGCGCGCGCTTCAAGCGCTG 638
 DB 651 CAGCGAGCTGG-----TGCAGCGCGAGCTGCGCGCGCGCTCTTCAACCGCGCT 701
 QY 639 CTGCGCGCGCAAGCAAGCAAGCGCGCAACTTTCAGCAACAAGAGCTGAGAAACA 698
 DB 702 GCGGCGCGCGCTGCGCTGAGTGAAGCG-----GATGAGCGCG 740
 QY 699 GGTGTGGGCAAGGTGGGCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 758
 DB 741 CTTGTGAGTGAAGCTTGTGCACTCAAGTGCAGGAGCGCAAGCGCGCGCGCGCG 800
 QY 759 CGCGCGCGCGCGCGCAATTCGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCTGTCGC 818
 DB 801 GCGTGGCGCGCGCGCGCGCAATTCGCGCAAGCTGCTCAATTCAGAGCAAGCGCGCGCG 860
 QY 819 GCGGCTACTGAGGCGCGCACTGCGCTTCAACGCAACGCGGCGCGCG---ACGTGAGCT 875
 DB 861 GCGGCACTGAGCGCGCGCACTGCGCTTCAACGCGCTTTCGCGCGCGCGCGCGCGCGCAT 920

QY 876 GTTGGCGGACCGCAAGTGGCTGTCGCAACCGCGCGCGCAAGCGCGCGCGCGCTGCTT 935
 DB 921 CTTGAGACAGAGAGAGAGTGGCGCGCTGAGAGCAAGCGCGCGCGCGCGCGCTGCTT 980
 QY 936 CTCCAGAGCGCGCTGCGCGCGCGCGAGAGCTTCTTGTGAGGTGGCGCTGCGCGCT 995
 DB 981 CACAGAGCGCGCTGAGAGGTGGCGAGAGCAATCTTCAATCAAGGTCAACGCGCTGCGCGCG 1040
 QY 996 GCGCGCGCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 1055
 DB 1041 GAGGAGAGAGAGAGAGAGCTGCTTGGAGGTCAACAGGTGACCTTGGAGCTGCGCGCG 1100
 QY 1056 CAAGAGTGTCCCGCGCAACCAAGAGCGCTGTCGACCGCAAGAGTACTGAGTGTGCG 1115
 DB 1101 CGGAGACTGCGCTTCAAGCCCGAGGCGCTGAGTGAACCGCAAGAGTTCGTGGCGCGTGG 1160
 QY 1116 GCGCGCGCGCGCGCGCTGCGCGAGCGCGCGAGCGCGCTCAAGCTTCAAGCTGCGCGCG 1175
 DB 1161 TCGCGTGGCGCGCGCTTGGCAAGCGCGCGCAATCTGCGCGCTGAGTGTCAACGCGAGCG 1220
 QY 1176 CGAGGTGCTCTGAGGATCAAGCGGAGTTCGCGCGCGCGCGCTGAGTGTGCAACAGCA 1235
 DB 1221 AAGAGCTGCACTGAGTCAACAGCGCGCGCGCGCGCGCGCGCGCGAGTGTGAGTGTCTG 1280
 QY 1236 GCAAGGCTGTGCGCGCTTCTTTCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCTGCT 1295
 DB 1281 GCAAGCGCTCTGAGAGTCTTCAAGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1340
 QY 1296 CGGTACCTTGAAGTCAAGCGCTGCGCAACAGACTCATAGAGTGTGCG----- 1342
 DB 1341 CTCACACATATGATGAGAACGAGGTGAGCCCATCTCCCTGCTCACTGCTTCCACTCC 1400
 QY 1343 -----TACAGGCGTCCCAAGAGATAGTATGATTAATGACCTTCAAGTGTCAACAGTCC 1397
 DB 1401 AACCTGACCAAGTGTGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1460
 QY 1398 CTCTCTGAGATCTGAGTCAATCCCTGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCTGCT 1457
 DB 1461 TGGAGCGCTGAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1520
 QY 1458 GTTCCCGGTGTTCTCCCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 1517
 DB 1521 GTCAACGAGTGAACCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1571
 QY 1518 CTTCGATGAGGAGTGAACAGGTGATCAATCAAGTGTGAGCAATGAGCTGTGAGCGCAAG 1577
 DB 1572 CTATGAAACAGGAGTGAATCAAGTCAATCAAGTGTGAGCAATGAGCTGTGAGTGTGAGTGTG 1631
 QY 1578 CTGCGCGCTGCGCGCTCAAGCGCAAGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCAT 1637
 DB 1632 CTGTGCGCTGCGCGCTCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 1691
 QY 1638 CAAGAGCTGATTAAGATCTAAG 1661
 DB 1692 CAAGAGCATATCAAGAGCTTACCG 1715

RESULT 13

ABX75855 standard, cDNA, 1725 BP.

XX ABX75855;

DT 31-MAR-2003 (first entry)

XX Rat Neu1 cDNA #1.

XX Human; mouse; rat; gene; 89; neutralized polypeptide; Neu; NHR; cancer;

KM neutralized homology repeat; C3H4 RING-zinc finger domain; myogenesis;

KM neurogenic function; transcriptional regulator; cell signalling; pain;

KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;

KM tumorigenesis; organ development; central nervous system; depression;

Db 1395 CACTCCAACTGCGCCAGTGCCTGGGAGACCGCCTCTCTGACCCCTGCTCAGACATG 1454
 Qy 1392 CCACTCTCTCTGCGCATCTGATGATCCCTGATGACGCCCCAGCTCCCGCTGAGCC 1451
 Db 1455 CCGCTTGGGCGCCCTGGGTGCTCTGTTGGCGGAACGCCCCCACTCACTGTGAGCCT 1514
 Qy 1452 CCGGTGTGTCCTGCTGTTCTCCCAACCGAGCCGCGCATCAAGATGAGCGAGCAG 1511
 Db 1515 GCCCGAGTCAACGATACCCCGGCTG-----GCCCATGAGAGCATGAATGAGCAG 1565
 Qy 1512 GGTGTCTTGCATGCGGAGTGAACACGTCATCTACAGTGTGACACATGTCCTGTG 1571
 Db 1566 CATTGCTATGAACACGCGATGATACATCTACAGCTGTGCGCACATGTGCTGTG 1625
 Qy 1572 CCACTGCTGCGCTGCGGCTCAAGCGACAGCCCGGCTGCTGCTGCTGCTGCGGCG 1631
 Db 1626 CTACTCTGTGGCTGCGCTCAAGAAAGCCCTGACGCTGTGCTGCTGCTGCGCTG 1685
 Qy 1632 GCCCATCAAGACGTCATTAAGATCTACAG 1661
 Db 1686 CCCCATCAAGACATCTATCAAGACTTACG 1715

RESULT 14

ABQ43990
ID ABQ43990 standard; DNA; 1016 BP.

AC ABQ43990;

DT 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 30581.

KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 KM SNP; cell differentiation; de.

OS Homo sapiens.

PN WO200218632-A2.

PD 07-MAR-2002.

PF 01-SEP-2001; 2001WO-BP10074.

PR 01-SEP-2000; 2000DE-1043826.

PR 05-SEP-2000; 2000DE-1044543.

PA (EPIC-) EPIDENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

DR WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA.

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (1) for diagnosis and/or prognosis of side effects of

CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (11) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

XX Sequence 1016 BP; 123 A; 154 C; 337 G; 399 T; 3 other;

Query Match 22.7%; Score 380.6; DB 24; Length 1016;

Best Local Similarity 72.6%; Pred. No. 56-52;

Matches 530; Conservative 0; Mismatches 197; Indels 3; Gaps 3;

Qy 571 CTTCTGGAAGAGGCGCTTGCTGCTGACAGCGTGAAGCCCGGCGCTGACGAGCCGCTTC 630
 Db 98 CGTATAGAGAGCGTTTCGTATATAG-TGACGTTCCGCGTTTATGTTAGTTTCGTTT 156
 Qy 631 AGCGCTGCTGCTGCGCCAGCAGCAGCAGCGCGCCACTTGCACAAACAGAGCTGAG 690
 Db 157 AGCGTTTGTTCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 216
 Qy 691 AACACAGAGTGTGCGCAAGCTGCGCACTGCGCTGCGCGCGCGCGCGCGCGCGCGCG 750
 Db 217 AATTATTAAGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 275
 Qy 751 CCAGCGGAGCG 810
 Db 276 TTATGACAGTGTGCG 335
 Qy 811 TCGTCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 870
 Db 336 TCGTCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 395
 Qy 871 AGCTGTGCG 930
 Db 396 AATTGTGTGCG 455
 Qy 931 GTCCTTCTCGAG-CG 989
 Db 456 GTTTTTCGAGATTTTCGTGCGTACGAGATTTTTCGTGAGAGTGTGCGTTC 515
 Qy 990 GGGGCTGCG 1049
 Db 516 GGGGTTGGCG 575
 Qy 1050 ACGGCGCAAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
 Db 576 ACGGTTTAAAGAGTGTGCGTCAATTAAACGCGTGTTCGATCTTAAAGAGTGTGG 635
 Qy 1110 GGTGGCG 1169
 Db 636 GGTGGCG 695
 Qy 1170 CCGCGCGCGAGTGTCTCTGCGCGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCG 1229
 Db 696 CCGCGCGCGAGTGTCTCTGCGCGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCG 755
 Qy 1230 CACGACGAGCGCTCTGCGCGCTTCTTCCGCGTGCAGCGCGCGCGCGCGCGCGCG 1289
 Db 756 TATTAGTATAGCGCTTTTGCGGTTTTCGTGCGCGCGCGCGCGCGCGCGCGCGCG 815
 Qy 1290 TCTCCGCGG 1299
 Db 816 TTTTTCGGT 825

RESULT 15

ABQ43991/C
ID ABQ43991 standard; DNA; 1016 BP.

XX

AC AB043991;
 XX 12-JUL-2002 (first entry)
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30582.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX SNP; cell differentiation; de.
 OS Homo sapiens.
 XX WO200218632-A2.
 PN 07-MAR-2002.
 PD 01-SEP-2001; 2001WO-EP10074.
 XX 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX (EPIC-) EPIDEMIOLOGICS AG.
 PA Olek A, Piepenbrock C, Berlin K, Guetig D;
 PI WPI; 2002-371829/40.
 DR Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised from the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (1) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (11) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 XX
 SQ Sequence 1016 BP; 399 A; 337 C; 154 G; 123 T; 3 other;

Query Match 22.7%; Score 380.6; DB 24; Length 1016;

Best Local Similarity 72.6%; Pred. No. 5e-52; Matches 530; Conservative 0; Mismatches 197; Indels 3; Gaps 3;

QY 571 CTTCTGAGAGCGCTTCGCTGACACGCTGACGCGCGCGCTTCAGCAGGCGCGCTTC 630
 DB 919 CGTATAGAGCGCTTCGCTGATAGC-TGACGCTTCGCGCTTCAGTTCGCTTC 861
 QY 631 AGCGCTGCTTCG 690
 DB 860 AGCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 801
 QY 691 AACACGACGAGCTGAGCG 750
 DB 800 AATTAATTAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742

QY 751 CCAAGCGACG 810
 DB 741 TTATCGACGTCG 682
 QY 811 TCGTCG 870
 DB 681 TCGTCG 622
 QY 871 AGCGTCG 930
 DB 621 AATTGTCG 562
 QY 931 GTCCTTCG 989
 DB 561 GTTTTTCGAGATTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 502
 QY 990 GCGGCTGCG 1049
 DB 501 GCGGCTGCG 442
 QY 1050 ACGGCG 1109
 DB 441 ACGGTTTAAAGATTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 382
 QY 1110 GGTGCG 1169
 DB 381 GGTGCG 322
 QY 1170 CGGCG 1229
 DB 321 CGGCG 262
 QY 1230 CACCAAGCG 1289
 DB 261 TATTACGTAAGCG 202
 QY 1290 TCT 1299
 DB 201 TTTTTCG 192

Search completed: January 20, 2004, 15:21:47
 Job time : 468 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 20, 2004, 17:03:09 ; Search time 571 Seconds
(without alignments)
10339.717 Million cell updates/sec

Title: US-09-808-387-21
Perfect score: 1675
Sequence: 1 atggcgacacagctgcacgc.....ctacagcgcatagctagcc 1675

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1675	100.0	1675	10	US-09-808-387-21
2	1228	73.3	1641	10	US-09-808-387-27
3	1103.6	65.9	1129	10	US-09-808-387-23
4	577.2	34.5	955	10	US-09-808-387-25
5	409.2	24.4	1674	10	US-09-808-387-1
6	409.2	24.4	1725	10	US-09-808-387-1
7	391.6	23.4	1698	10	US-09-808-387-9
8	391.6	23.4	1749	10	US-09-808-387-7
9	380.8	22.7	1725	10	US-09-808-387-15
10	264	15.8	1161	10	US-09-808-387-5
11	186.6	11.1	694	13	US-10-029-386-22715
12	182.4	10.9	876	10	US-09-808-387-13
13	177	10.6	888	10	US-09-808-387-19
14	176.8	10.6	1035	10	US-09-808-387-17
15	176.4	10.5	1035	10	US-09-808-387-11

16	166.4	9.9	789	10	US-09-808-387-29	Sequence 29, Appl
17	135.8	8.1	765	10	US-09-808-387-31	Sequence 31, Appl
18	127.8	7.6	693	13	US-10-027-633-152858	Sequence 152858,
19	127.8	7.6	693	14	US-10-027-633-152858	Sequence 152858,
20	119	7.1	536	13	US-10-029-386-5702	Sequence 5702, Ap
21	118	7.0	523	13	US-10-029-386-19512	Sequence 19512, A
22	109.8	6.6	523	13	US-10-029-386-8783	Sequence 8783, Ap
23	108.2	6.5	239	13	US-10-029-386-22509	Sequence 22509, A
24	93	5.6	2307	11	US-09-893-519A-87	Sequence 87, Appl
25	84.8	5.1	2886	15	US-10-156-761-1966	Sequence 1966, Ap
26	84.8	5.1	9025608	15	US-10-156-761-1	Sequence 1, Appl
27	83.4	5.0	3957	13	US-10-200-562-193	Sequence 193, App
28	83.4	5.0	3957	13	US-10-237-551-193	Sequence 193, App
29	83.4	5.0	154746	13	US-09-827-688-8	Sequence 8, Appl
30	83.4	5.0	154746	13	US-09-827-688-8	Sequence 8, Appl
31	81.8	4.9	125746	15	US-10-156-761-15102	Sequence 15102, A
32	81.8	4.9	9025608	15	US-10-156-761-1	Sequence 1, Appl
33	79.4	4.7	1392	13	US-10-214-446-7	Sequence 7, Appl
34	79.4	4.7	4689	10	US-09-861-289-34	Sequence 34, Appl
35	79.4	4.7	4689	10	US-09-860-846-34	Sequence 34, Appl
36	79.4	4.7	4689	11	US-09-988-384B-34	Sequence 34, Appl
37	79.4	4.7	4689	11	US-09-988-384B-34	Sequence 34, Appl
38	79.4	4.7	4689	13	US-10-271-889-34	Sequence 34, Appl
39	79.4	4.7	36778	10	US-09-861-289-5	Sequence 5, Appl
40	79.4	4.7	36778	10	US-09-860-846-5	Sequence 5, Appl
41	79.4	4.7	36778	11	US-09-836-821-5	Sequence 5, Appl
42	79.4	4.7	36778	13	US-10-271-889-48	Sequence 48, Appl
43	79.4	4.7	37948	11	US-09-988-384B-5	Sequence 5, Appl
44	79.4	4.7	38506	11	US-09-793-708-19	Sequence 19, Appl
45	79.4	4.7	38506	13	US-10-201-365-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-808-387-21
Sequence 21, Application US/09808387
Patent No. US20020132293A1
GENERAL INFORMATION:
APPLICANT: Kaia Palm
APPLICANT: Tonie Timmek
TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
FILE REFERENCE: CEMRES.001A
CURRENT APPLICATION NUMBER: US/09/808,387
CURRENT FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 1675
TYPE: DNA
ORGANISM: Homo sapien
US-09-808-387-21

Query Match 100.0%; Score 1675; DB 10; Length 1675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGCAACACGGTGACCCGACCTTCCAGACCCGAGCCACCGCGCTCTGGCC	60
DB	1	ATGGGCAACACGGTGACCCGACCTTCCAGACCCGAGCCACCGCGCTCTGGCC	60
QY	61	ACCGGCGGCTGCGGCGCCCGGCGGAGGACGCGCGATCTGGGCGAGCGCGCGC	120
DB	61	ACCGGCGGCTGCGGCGCCCGGCGGAGGACGCGCGATCTGGGCGAGCGCGCGC	120
QY	121	TTCACCGGACGAGGCAAGGCAAGACGTCGCGCTGACGCGCACTCGCGCGGCGACA	180
DB	121	TTCACCGGACGAGGCAAGGCAAGACGTCGCGCTGACGCGCACTCGCGCGGCGACA	180
QY	181	CGGCGCAACAGCTTGCAGATGGCGTCACTTACGACGAGCGCCCATCCGGCTGTACAG	240

Db	181	CGGCGCAACACTTCTGCAATGAGGTCAAGTTCAAGCAAGCGGCGCCATCCGACTTACGAG	240
Qy	241	CAGGTGGGCGCTGCGCCTTGATGAGCGTGGCGCCTTGAGTGAAGCGGCGCTGCGCTTGGCG	300
Db	241	CAGGTGGGCGCTGCGCCTTGATGAGCGTGGCGCCTTGAGTGAAGCGGCGCTGCGCTTGGCG	300
Qy	301	TTCAACCGGCAAGATTCGTCGCTCAATGAGCGCCCAAGACATCCCAAGTACGCTTGGCG	360
Db	301	TTCAACCGGCAAGATTCGTCGCTCAATGAGCGCCCAAGACATCCCAAGTACGCTTGGCG	360
Qy	361	GACCTGGTCAAGCGGCGCGGAGCTACGAGGCGCAAGGCACTGGCCCGAAGAACCTGAGCGCTGGC	420
Db	361	GACCTGGTCAAGCGGCGCGGAGCTACGAGGCGCAAGGCACTGGCCCGAAGAACCTGAGCGCTGGC	420
Qy	421	GACACGAGTGTGCGCTTACGAGGCGGACCGGCAAGCGCGCGTGTCTACAGCGTGAACGAC	480
Db	421	GACACGAGTGTGCGCTTACGAGGCGGACCGGCAAGCGCGCGTGTCTACAGCGTGAACGAC	480
Qy	481	GCGAGCGGAGTCTTCTTTCACATGCGGCGTGGCCGTGGCGGCGCGCTCTGGGCGCTCAT	540
Db	481	GCGAGCGGAGTCTTCTTTCACATGCGGCGTGGCCGTGGCGGCGCGCTCTGGGCGCTCAT	540
Qy	541	GATGTCTACGGGATACCGACGAGAGTGAAGCTTCTGGAAGACCGCTTGGCTGAACCGTG	600
Db	541	GATGTCTACGGGATACCGACGAGAGTGAAGCTTCTGGAAGACCGCTTGGCTGAACCGTG	600
Qy	601	ACGCGCGGCGGCTTCAAGCAGGAGCGCGCTTCAAGCGCGCTGTCGCGCGCCAGACGACGAC	660
Db	601	ACGCGCGGCGGCTTCAAGCAGGAGCGCGCTTCAAGCGCGCTGTCGCGCGCCAGACGACGAC	660
Qy	661	GCGGCGCACTTCAGCAACAGAGTTCGAGAACAAACAGAGTGTGGCCAGACTGAGCCAC	720
Db	661	GCGGCGCACTTCAGCAACAGAGTTCGAGAACAAACAGAGTGTGGCCAGACTGAGCCAC	720
Qy	721	CTGGGCGTGGGCGCGGCGCCCGGAGCCCAACCGGCAAGCGGCGCGCGCGCGCATTCGG	780
Db	721	CTGGGCGTGGGCGCGGCGCCCGGAGCCCAACCGGCAAGCGGCGCGCGCGCGCATTCGG	780
Qy	781	TGCGGAGGCGCGTGAAGCGCGCGGCGCGCGGTCGTGCGCGGCGCTTACGAGAGCGGACCTG	840
Db	781	TGCGGAGGCGCGTGAAGCGCGCGGCGCGCGGTCGTGCGCGGCGCTTACGAGAGCGGACCTG	840
Qy	841	CGCTTTCACGCAACACGCGGAGCCGACGTTGAGCGCTTGGCGCGAACCGCAAAAGTGGCTGC	900
Db	841	CGCTTTCACGCAACACGCGGAGCCGACGTTGAGCGCTTGGCGCGAACCGCAAAAGTGGCTGC	900
Qy	901	GCAACCGGAGCCCGACGAGCGCGCGCGCGCTGCTTCTTCGAGAGCGCGCTGGCGGCGG	960
Db	901	GCAACCGGAGCCCGACGAGCGCGCGCGCGCTGCTTCTTCGAGAGCGCGCTGGCGGCGG	960
Qy	961	GAGAGCGCTTTCGTGGAAGTGAAGTGGGCGGTCGAGGAGCTGAGGCGCGCGCGCGCTTTC	1020
Db	961	GAGAGCGCTTTCGTGGAAGTGAAGTGGGCGGTCGAGGAGCTGAGGCGCGCGCGCGCTTTC	1020
Qy	1021	GGCATCACAGTCAAGTGAAGCCCGGAGCGTGTACAGGCGGCTTCCAGAGCTGCGCGGAC	1080
Db	1021	GGCATCACAGTCAAGTGAAGCCCGGAGCGTGTACAGGCGGCTTCCAGAGCTGCGCGGAC	1080
Qy	1081	GCGCTGCTCGAACCGCAAAAGTACTGGAGTGGGCGCGCGCGGAGCGCGCTGCGGAGCGG	1140
Db	1081	GCGCTGCTCGAACCGCAAAAGTACTGGAGTGGGCGCGCGCGGAGCGCGCTGCGGAGCGG	1140
Qy	1141	GCGCAACGCGCTCAAGTTCACGCTGAGCGCGCGCGCGCGGACGTCCTCTGAGGACATCAACGG	1200
Db	1141	GCGCAACGCGCTCAAGTTCACGCTGAGCGCGCGCGCGCGGACGTCCTCTGAGGACATCAACGG	1200
Qy	1201	GCTCCGCGCGGCGCGCTGCTGTCGCTGACACACAGCAGGCGGCTCTGAGGCGCTTCTTCCGC	1260
Db	1201	GCTCCGCGCGGCGCGCTGCTGTCGCTGACACACAGCAGGCGGCTCTGAGGCGCTTCTTCCGC	1260
Qy	1261	GTGCGCGGCGGCTGCGGAGCGCAAGCTGCGCTCTCTCGAGTACCCTTGACATCGACCTTGGC	1320

[illegible]

Db 301 TTGCTGCGACGACCGGTGCTCATGAGCGCACAGATATCCCAAGTACGCTGCGCC 360
Qy 361 GACCTGGTACAGCGGCGGCTACTGAGGCGAAGGCACTGCGCGAAGAACTGGCGTGGC 420
Db 361 GACCTGGTACACAGACTGATCTGGGCGAAGGCGCTGGCGGAGAACCTGGCGTGGC 420
Qy 421 GACAGGCTGCTGAGCTACTGAGGCGGACCGCGCGCGGTCTTCAAGCGTGAACGAC 480
Db 421 GACAGGCTGCTGAGCTACTGAGGCGGACCGCGCGCGGTCTTCAAGCGTGAACGAC 480
Qy 481 GAGCGAGCGGCTGCTTCTCACTGGCGGCTGGCCGTGGCGGCGCGCTTGGCGCTCAT 540
Db 481 GAGCGAGCGGCTGCTTCTCACTGGCGGCTGGCCGTGGCGGCGCGCTTGGCGCTCAT 540
Qy 481 GAGCGAGCGGCTGCTTCTCACTGGCGGCTGGCCGTGGCGGCGCGCTTGGCGCTCAT 540
Db 481 GAGCGAGCGGCTGCTTCTCACTGGCGGCTGGCCGTGGCGGCGCGCTTGGCGCTCAT 540
Qy 541 GATGCTTACCGGATCACCGAGAGGTGAGCTTCTGAGAGCGCGCTTGGCTGACAGCTG 600
Db 541 GATGCTTACCGGATCACCGAGAGGTGAGCTTCTGAGAGCGCGCTTGGCTGACAGCTG 600
Qy 541 GATGCTTACCGGATCACCGAGAGGTGAGCTTCTGAGAGCGCGCTTGGCTGACAGCTG 600
Db 541 GATGCTTACCGGATCACCGAGAGGTGAGCTTCTGAGAGCGCGCTTGGCTGACAGCTG 600
Qy 601 ACGCCGCGCGCTTCAAGCAGGCGCGCTTCAAGCAGGCGCGCTTCAAGCAGGCGCGCT 660
Db 601 ACGCCGCGCGCTTCAAGCAGGCGCGCTTCAAGCAGGCGCGCTTCAAGCAGGCGCGCT 660
Qy 661 GCGGCGCACTTGCACACAGAGCTTGCAGAGAACACAGAGGTGGGCGCAAGCTGGCGAC 720
Db 661 GCGGCGCACTTGCACACAGAGCTTGCAGAGAACACAGAGGTGGGCGCAAGCTGGCGAC 720
Qy 661 GCTGCGCACTTGCATATATACAGAGCTGAGAGATACAGAGGTGGGCGCAAGCTGGCGAC 720
Db 661 GCTGCGCACTTGCATATATACAGAGCTGAGAGATACAGAGGTGGGCGCAAGCTGGCGAC 720
Qy 721 CTGGCGGCTGGCG 780
Db 721 CTGGCGGCTGGCG 780
Qy 721 CTGGCGGCTGGCG 780
Db 721 CTGGCGGCTGGCG 780
Qy 781 TGCGGCG 840
Db 781 TGCGGCG 840
Qy 754 TGCGGCG 813
Db 754 TGCGGCG 813
Qy 841 CGCTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 CGCTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 814 CGTTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 873
Db 814 CGTTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 873
Qy 901 GCACCGGCG 960
Db 901 GCACCGGCG 960
Qy 874 GCG 933
Db 874 GCG 933
Qy 961 GAGAGCGCTTCTGAGAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 961 GAGAGCGCTTCTGAGAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Qy 934 GAGAGCGCTTCTGAGAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993
Db 934 GAGAGCGCTTCTGAGAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993
Qy 1021 GGCATCACGCTGCGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 1021 GGCATCACGCTGCGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy 994 GGCATCACGCTGCGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1053
Db 994 GGCATCACGCTGCGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1053
Qy 1081 GCGGCTGCTGAGCG 1140
Db 1081 GCGGCTGCTGAGCG 1140
Qy 1054 GCGGCTGCTGAGCG 1113
Db 1054 GCGGCTGCTGAGCG 1113
Qy 1141 GCGGCGCGCTCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1200
Db 1141 GCGGCGCGCTCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1200
Qy 1114 GCGGCGCGCTCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1173
Db 1114 GCGGCGCGCTCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1173
Qy 1201 GCGGCG 1260
Db 1201 GCGGCG 1260
Qy 1174 GCGGCG 1233
Db 1174 GCGGCG 1233
Qy 1261 GCGGCG 1320
Db 1261 GCGGCG 1320
Qy 1234 GCGGCG 1293
Db 1234 GCGGCG 1293
Qy 1321 ACACGAGCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAG 1380
Db 1321 ACACGAGCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAG 1380
Qy 1294 GCGGCG 1353
Db 1294 GCGGCG 1353
Qy 1381 TTGAGTGAAGGAGTCTCTCTGCGGATCTGAGTATCTCTGAGTATCTCTGAGTATCTCT 1440
Db 1381 TTGAGTGAAGGAGTCTCTCTGCGGATCTGAGTATCTCTGAGTATCTCTGAGTATCTCT 1440

Db 1354 TTGAGGAGTCAACAGTGTATCATGAGTATCAAGTATCTCTGAGTATCTCTGAGTATCTCT 1413
Qy 1441 CCGGTGAGCG 1500
Db 1414 CCACTGAGTCTCCCGAGTGTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1473
Qy 1501 GCGGAGTGAAGGAGTGTCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560
Db 1474 GAGAGTGAAGGAGTGTCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1533
Qy 1561 ATGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1620
Db 1534 ATGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1593
Qy 1621 ATGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1688
Db 1594 ATGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1641

RESULT 3
US-09-808-387-23
Sequence 23, Application US/09808387
Patent No. US20020132293A1
GENERAL INFORMATION:
APPLICANT: Kala Palm
APPLICANT: Tom's Timmusk
APPLICANT: Cemines Research
TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
FILE REFERENCE: CEMRES 001A
CURRENT APPLICATION NUMBER: US/09/808,387
CURRENT FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 1129
TYPE: DNA
ORGANISM: Homo sapien
US-09-808-387-23

Query Match 65.9%; Score 1103.6; DB 10; Length 1129;
Best Local Similarity 98.8%; Pred. No. 3,1e-235;
Matches 1112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 550 GGCATCACGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 609
Db 4 GGCATCACGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 63
Qy 610 GCGCTTCAAGCG 669
Db 64 GCGCTTCAAGCG 123
Qy 670 TTGAGCAACAGAGCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 729
Db 124 TTGAGCAACAGAGCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 183
Qy 730 GCG 789
Db 184 GCG 243
Qy 790 GCGGAGCG 849
Db 244 GCGGAGCG 303
Qy 850 GCAACAGCG 909
Db 304 GCAACAGCG 363
Qy 910 CCGGAGCG 969
Db 364 CCGGAGCG 423
Qy 970 TTGAGTGAAGGAGTCTCTCTGCGGATCTGAGTATCTCTGAGTATCTCTGAGTATCTCT 1029
Db 970 TTGAGTGAAGGAGTCTCTCTGCGGATCTGAGTATCTCTGAGTATCTCTGAGTATCTCT 1029

Db 424 TTGTTGAGAGGTGGCCGCTGCGGGGCTGGCGCGCCCGCGCTGGCTTGCGGCATCAACG 483
1030 TCGTGGAGACCGGGGCGGTCTACAGGCCCAAGAGCTGGCCGCGACCAAGACGGGCTGTC 1089
484 TCGTGGAGACCGGGGCGGTCTACAGGCCCAAGAGCTGGCCGCGACCAAGACGGGCTGTC 543
Qy 1090 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1149
544 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603
Qy 1150 CTCAGCTTCAAGCTGCG 1209
604 CTCAGCTTCAAGCTGCG 663
Qy 1210 GGGCGGCTGCTGGTGGCG 1269
664 GGGCGGCTGCTGGTGGCG 723
Db 1270 GGGCGGCTGCTGGTGGCG 1329
724 GGGCGGCTGCTGGTGGCG 783
Qy 1330 CCATCAAGAGTCTCCAGCGCGCTCCAGAGCATATGATTAAGATTAAGATTAAGATTA 1389
784 CCATCAAGAGTCTCCAGCGCGCTCCAGAGCATATGATTAAGATTAAGATTAAGATTA 843
Qy 1390 AACAGTCTCTCTGGGATCTGAGTCAATCCCTGGTGAAGCGCGCGCGCGCGCGCGCG 1449
844 AACAGTCTCTCTGGGATCTGAGTCAATCCCTGGTGAAGCGCGCGCGCGCGCGCGCG 903
Qy 1450 CCCCCGCTGCTCCCGCTGTTCTCCCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509
904 CCCCCGCTGCTCCCGCTGTTCTCCCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 963
Qy 1510 ACGGTGTCTTCAATGGCGAGGTGAGACAGGTCACTTCAACGTTGTGACACATGTGCTG 1569
964 ACGGTGTCTTCAATGGCGAGGTGAGACAGGTCACTTCAACGTTGTGACACATGTGCTG 1023
Db 1570 TGGCAGAGCTGGCGCGCGCGCGCTCAAGCGACGCGCGCGCGCGCGCGCGCGCGCG 1629
1024 TGGCAGAGCTGGCGCGCGCGCGCTCAAGCGACGCGCGCGCGCGCGCGCGCGCGCG 1083
Qy 1630 CGGCGCATCAAGAGCTCATTAAGATTAAGAGCGCATAGCTAGCC 1675
1084 CGGCGCATCAAGAGCTCATTAAGATTAAGAGCGCATAGCTAGCC 1129
Db

RESULT 4
US-09-808-387-25
; Sequence 25, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-25

Query Match 34.5%; Score 577.2; DB 10; Length 955;
Best Local Similarity 99.5%; Pred. No. 9.9e-119;
Matches 579; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCAACAGGTGACCGGACCTTGGCAAGACCGGACCGGCGCGCTCTGCGCC 60
Db 1 ATGGGCAACAGGTGACCGGACCTTGGCAAGACCGGACCGGCGCGCTCTGCGCC 60
Qy 61 ACCCGGCGGTGCTGGCG 120
61 ACCCGGCGGTGCTGGCG 120
Db 61 ACCCGGCGGTGCTGGCG 120
Qy 121 TTCCAGCGCAGGCGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCA 180
121 TTCCAGCGCAGGCGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCA 180
Db 121 TTCCAGCGCAGGCGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCAAGAGGCA 180
Qy 181 CCGCGCAAGAGCTTCTGCAATGAGGCTCAAGCTTCAAGAGGCGCGCGCGCGCGCGCG 240
181 CCGCGCAAGAGCTTCTGCAATGAGGCTCAAGCTTCAAGAGGCGCGCGCGCGCGCGCG 240
Db 181 CCGCGCAAGAGCTTCTGCAATGAGGCTCAAGCTTCAAGAGGCGCGCGCGCGCGCGCG 240
Qy 241 CAGGTGGGCTGGCG 300
241 CAGGTGGGCTGGCG 300
Db 241 CAGGTGGGCTGGCG 300
Qy 301 TTCAACGCGCAGATCCGTGCTCATAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
301 TTCAACGCGCAGATCCGTGCTCATAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 TTCAACGCGCAGATCCGTGCTCATAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 GACCTGTCAAGCG 420
361 GACCTGTCAAGCG 420
Db 361 GACCTGTCAAGCG 420
Qy 421 GACAGGCTGCTGGCG 480
421 GACAGGCTGCTGGCG 480
Db 421 GACAGGCTGCTGGCG 480
Qy 481 GCGGAGCGGTGCTTTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
481 GCGGAGCGGTGCTTTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 GCGGAGCGGTGCTTTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy 541 GATGTCTAAGCGCATCAACGAGGTGCAAGCTTCTGAGAGC 582
541 GATGTCTAAGCGCATCAACGAGGTGCAAGCTTCTGAGAGC 582
Db 541 GATGTCTAAGCGCATCAACGAGGTGCAAGCTTCTGAGAGC 582

RESULT 5
US-09-808-387-3
; Sequence 3, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-3

Query Match 24.4%; Score 409.2; DB 10; Length 1674;
Best Local Similarity 56.6%; Pred. No. 1.4e-81;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

Qy 99 GGTCTGGGCGGAGCGCGCGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 158
Db 120 GCTCCAGCGCACCGCGCTGCTTTCACCGCGCACAGCGCGCTCCCAATCTCATGGA 179
Qy 159 CGGCGACTCG 218
159 CGGCGACTCG 218
Db 180 CTTGAGCGCACAGGCTGTCAAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239

QY 219 GCGGCCATCGGCTGTACAGAGAGGTCGGCTGCGCTGTGGCCCTGGCTG 278
 DB 240 CCGCCCGGCTCTCATCTACAGAGAGGTCAGGCTAGAGATCAACAAAGAGAGTGCCTG 299
 QY 279 GAGGAGGAGGCTGTGGCTTACCGGCGACAGATCCGCTGTCAATGAGCGCCAGGA 338
 DB 300 GAGGAGGAGGCTGTGGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
 QY 339 CATCCCAAGTACGCTGCGGACCTGTGTACCGGCGCGGCTTACTGAGCCAGAGGACT 398
 DB 360 GGTGCCAAGTACGCTGCGGACCTGTGTACCGAGAGTGTCTTGTGGGCAAGGCGCT 419
 QY 399 GCGGAGAGAGCTGGCGCTGCGGACAGAGTGTGGCTTGTGGGCGACGCGCGCG 458
 DB 420 GCTGAGAGAGTGTGCTCAATGAGGCAATCATGTGTGTGTGTGTGTGTGTGTGTGT 479
 QY 459 GGT 518
 DB 480 TGTCTTCAACGATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539
 QY 519 GCGCCGCTGTGGCGCTCATTTGATGTCTACGAGATCAACGAGAGTGTGACTTGTGA 578
 DB 540 GAGCCGCTGTGGCGCTGA 599
 QY 579 GAGGCGCTGT 638
 DB 600 TAGGAGAGT 650
 QY 639 CTGTGCGGCGGAG 698
 DB 651 GCGGCGGCGGCTGT 699
 QY 699 GATGT 758
 DB 700 -----AGCTTATGT 749
 QY 759 GCGGCGGCGGCGGCTGT 818
 DB 750 GCGGCGGCGGCTGT 809
 QY 819 GCGGCTTGT 875
 DB 810 GCGGCGGCTGT 869
 QY 876 GTGCGGCGGAG 935
 DB 870 CTGTGAG 929
 QY 936 CTGTGAG 995
 DB 930 CACAG 989
 QY 996 GCGGCGGCGGCGGCTGT 1055
 DB 990 GCGGCGGCGGCGGCTGT 1049
 QY 1056 CAAG 1115
 DB 1050 GCGGCGGCTGT 1109
 QY 1116 GCGGCGGCGGCGGCTGT 1175
 DB 1110 CCGGCGGCGGCGGCTGT 1169
 QY 1176 GAGAGT 1235
 DB 1170 GAGAGT 1229
 QY 1236 GAG 1295
 DB 1230 GAG 1289

QY 1296 GAGTACCTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1346
 DB 1290 CTGCACTATCTGTGAG 1349
 QY 1347 -----GAGTGTGAG 1397
 DB 1350 AAGTGTGAG 1409
 QY 1398 CTGCTGTGAG 1457
 DB 1410 TGTGCTGT 1469
 QY 1458 GTGCGGCGGCTGT 1517
 DB 1470 GTGCGGCGGAG 1520
 QY 1518 CTGT 1577
 DB 1521 CTATGT 1580
 QY 1578 CTGT 1637
 DB 1581 CTGT 1640
 QY 1638 CAAGAGAGT 1661
 DB 1641 CAAGAGAGT 1664

RESULT 6
 US-09-808-387-1
 ; Sequence 1, Application US/09808387
 ; Patent No. US20020132293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaia Palm
 ; APPLICANT: Tonis Timmusk
 ; APPLICANT: Cemines Research
 ; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
 ; FILE REFERENCE: CEMRES.001A
 ; CURRENT APPLICATION NUMBER: US/09/808,387
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatsSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1725
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-808-387-1
 Query Match 24.4%, Score 409.2; DB 10; Length 1725;
 Best Local Similarity 56.6%; Pred. No. 1,4e-81;
 Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;
 QY 99 GATCTGT 158
 DB 171 GCTGT 230
 QY 159 GCGGCACTGT 218
 DB 231 CTGTGAGCAAGAGCTGT 290
 QY 219 GCGGCGGCTGT 278
 DB 291 CCGGCGGCTGT 350
 QY 279 GAGGAGGAG 338
 DB 351 GAGGAGGAG 410
 QY 339 CATCCCAAGTACGCTGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 398
 DB 411 GCTGCCAAGTACGCTGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470

QY 399 GCCCGAAGACTGCGCTGCGGACACAGGCTGCTGCTACTGAGCCGACCGGACGCGCG 458
 DB 471 GCTTGAAGAGTTTGGCCATGAGGAGGACATCATTCGATTCCTGGGTGACAAAGAGCGCG 530
 QY 459 CGTGTCTTACAGCGTGAACGACGCGGAGCGGCTCTTTCATCTGCGGCGTGGCGCTGGG 518
 DB 531 TGTCTTCCACCGCATCAACGACTCGCTGTATGCTGTCTTTCAGCGGAGGTCCGACGCGC 590
 QY 519 CGGCGGCTCTGAGGCGCTATGATGTCTAACCGGATCAACGAGAGGTGACGTTTGA 578
 DB 591 CACACCGCTCTGGGCGCTGTGACGCTTACCGGCTCACGCGGAGGTCAAGCTGCTTGA 650
 QY 579 GAGCGGCTTCCGCTGACAGCTGACGCGCGCGCTCACGAGGCGCGCTTACGCGCTG 638
 DB 651 TAGCGAGCTGG-----TGTCTCCGACCTGTCTGCGGCGCGCTCTTCAACCGCT 701
 QY 639 CTTGCGGCGCAGACGACGACGCGGCGCACTTTCGACAAACAGAGCTGAGAGAACCA 698
 DB 702 GCGGCGGCGCTGCTGCTGCGGCGGAGGCGGACGACGCGCGCTCTCGGTG----- 750
 QY 699 GGTGTGCGCAAGCTGAGGCGCACTTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCG 758
 DB 751 -----AGCTATGCGACTCAACGCTGCGGCGCGGACGCGGCGCGCGCGCGCGCG 800
 QY 759 CGCGCGGCGCGCGCGCATTCCTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818
 DB 801 GCGCGCGGCTGCGCGCATCCCGAGAACTCACTCACTCGGACGACGCGCGCGCGCGCG 860
 QY 819 GCGGCTACTGAGAGCGGACCTGCGCTTTCAGCGCAACGCGGCGCGCGCGCGCGCGCG 875
 DB 861 GCGGCGAGCTGACGCGGACCTGCGCTTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 920
 QY 876 GTCGCGCGACCGGAAAGTGGCTTGGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCG 935
 DB 921 CCTCGAGAGACAGACGCGTGGCGCGCTGAGGACGCGGCGCGGACGCGCGCTCGCTT 980
 QY 936 CTTGCGAGCGCGCGCTGCGCGCGCGCGCGGAGGCTCTTCTGTGAGAGTGGCGCGTGG 995
 DB 981 CACGACCGCGCGCGCTGCGCGCGCGCGCGGACGATCTTCTGTCAAGTCAACGCGCTG 1040
 QY 996 GCGGCGCGCGCGCGCGCTTTCGCGCATCAGCTGTGCGGACCGCGGCGGTCTACGCG 1055
 DB 1041 CGCGCGCGCGCGCGCGCTGCTGCTGCGCGCTCACGAGTGGACCGCGCGCGCGCGCG 1100
 QY 1056 CAACGAGCTGCGCGCGCGGACCGGACGCGCTGCGGACCGGAGAGAGTCTGGGTGG 1115
 DB 1101 GCGCGACCTGCGCTTTCAGCGCTTGAAGCGCTGTGAGACCGGAGAAATTCGGGCG 1160
 QY 1116 GCGCGCGCGCGCGCGCTGCGGAGCGCGCGGACGCGCTCAGCTTACGCTGCGCGCG 1175
 DB 1161 CGCGCGCGCGCGCGCGCTGCGGACGCGGACGATCTGTGGGCTGTGTCAACCGGAC 1220
 QY 1176 CGAGCTGCTTCTGGGACATCAACGCGCGCTGCGCGCGCGCGCTGTGTGTCAACG 1235
 DB 1221 CGAGCTGACACTCAGGCAACATGCGCGCGCGCGCGGACGAGCTGTGACTGTGAC 1280
 QY 1236 GAGGCGGCTCTGCGCGCTTCTTCCCGCGTGGCGCGCGCGCGGTGCGCGGCGAGTGT 1295
 DB 1281 GACGCGGCTTGTGATGCTTTCGCGCTGCGGACGCGGACATGACGAGTCTCGCAT 1340
 QY 1296 CGGTACCTTGAAGTCAAGCGCTGCGGACGAGCTCCATCAAGGTCTCTCAG----- 1346
 DB 1341 CTTCATATCTTGGCGGAGCGGAGGTATCCGTGATCTCCCTGCTCCCTGCGCGCG 1400
 QY 1347 -----CGGCTCCGAGAGCATATGATTTCAATATGACCTTCAATGTCAACGAGTC 1397
 DB 1401 AACCTGCGGAGGCGCTGCGGAGCGCGCTGTGACCGCTTGTGAGCAAGTCAAGCTC 1460
 QY 1398 CTCTGCGGAGTCAAGTCACTCGGTGAGAGCGCGCGCGCGCGCTCCCGGTGAGCGCG 1457
 DB 1461 TGGCGCTTGTGGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1520

QY 1458 GTCCCCCGTGTCTTCCCAACCGGAGCGGACAGGATCAAGAAAGGAGTGCAGGTGTG 1517
 DB 1521 GTGCGCAATGACCCCAAGGTCTGGGCGAGTG-----GAGCGATGAGTGACCATTTG 1571
 QY 1518 CTTGAGATGGAGGTGAGCAGGTGTATCTACACGATGAGGACATATGCTGTGCGACAG 1577
 DB 1572 CTATGAAACAGCGGTGAGACAGGTATCTACATATGAGGACATATGCTGTGCGACG 1631
 QY 1578 CTGCGGCTGCGGCTCAAGCGGACAGGCGCGGCTGTGCGCGCATCTGCGCGCGCG 1637
 DB 1632 CTGTGCGCTGCGGCTCAAGGAGGCTGTGACGCGCTGTGCGCGCATCTGCGCGCG 1691
 QY 1638 CAAGAGCGTCAATTAAGTCTACAG 1661
 DB 1692 CAAGGACATCATCAAGACTACG 1715

 RESULT 7
 US-09-808-387-9
 ; Sequence 9, Application US/09808387
 ; Patent No. US20020132293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaja Palm
 ; APPLICANT: Tonis Timmusk
 ; APPLICANT: Cemines Research
 ; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
 ; FILE REFERENCE: TRANSRIPTION REGULATORS AND USHS THEREFOR
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSB for Windows Version 4.0
 ; SEQ ID NO 9
 ; TYPE: DNA
 ; ORGANISM: mouse
 US-09-808-387-9

 Query Match 23.4%; Score 391.6; DB 10; Length 1698;
 Best Local Similarity 55.9%; Pred. No. 1.1e-77;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

 QY 99 GGTCTGTGGAGGCGCGCGCTTTCAGCGCGGACGCGGCAAGAGGAGTGTGCGCTGGA 158
 DB 120 GCTCCCGGACAGCGCGCTGCTCTTCCACCCGACATAAGGCGCTCCAGATCTCATGGA 179
 QY 159 CGGCACTCGCGCGCGGCGGACAGCGGCGGACAGCTTCTGCAATGAGGTGACGCTACGGA 218
 DB 180 CTTGAGCGCAAGGCGCGCTGACAGAGGAGGCGGAGCTTGTGCAATGAGCTTCACTGTA 239
 QY 219 GCGGCGCATCGGCTGTATGAGCAGAGTGGCGGCTGCGGCTGTGTGCGCGCGCGCTG 278
 DB 240 CGCGCGGCTGCTATCTATGAGGAGGAGTCAAGGCTGTGAAATCAACAGAGGAGATGCTG 299
 QY 279 GAGCGGCGGCTGCGCTTCTGCTTACCGGCGGACGATCCGTGTGCTATAGGCGCGGAG 338
 DB 300 GAGCGGCGGCTGCGGCTTCTGCTTCAACGAGAGGACCTTCCGATTCACCGCGACTC 359
 QY 339 CATCCCAAGTACGCTGCGGAGCGGAGCTGTGACGCGGCGGAGCTATGCGGCGGAGCACT 398
 DB 360 GCTGCGCAAGTACGCTGCGGAGCTGTGAGTGTCTGAGATGAGCTTCTGAGGCGGAGGAT 419
 QY 399 GCCCGAAGACTTGGCGCTGCGGAGCAGGTGTGCTTACTTGGCGGAGCGGCGGCGG 458
 DB 420 GCTGAGAGGTTTGGCAAGAGGAGGCAATCATTTGCTTCTGAGGTGAGCAAGAGGCGG 479
 QY 459 CGTGTCTTACAGCGTGAACGAGGAGGCGGCTTCTTCACTGCGGCGGTGCGCGGTGG 518
 DB 480 CGTCTTCTACCGGATCAAGATGAGCTGTATGCTTCTTCTTCAAGTGGGAGTCCGAGCGGT 539
 QY 519 CGGCGCGCTGCGGCGGCTCATTTGATGTCTACGCGGATCAACGAGAGGTGACGCTTGTGA 578
 DB 540 GAGCGGCTCTGGGCGCTGTGTGAGAGTCTACGCGGCTCACGCGGAGGTGTCTACGCTTGA 599

QY 579 GAGGCGCTTTCGTGACAGCTGACGCGCGCGCTTCAAGCAGCGCGCTTCAAGCGCTG 638
 DB 600 CAGCGAGCTGG-----TGTGCGCGAGCTGCGCGCGCGCTTCAAGCGCGCT 650
 QY 639 CCGTCCCGCGCAGAGCAGACAGCGCGCGCACTTGCAGAACAAAGAGCTTGAAGAACCA 698
 DB 651 GCGGCGCGCTGCGCTGCGCGTGAAGCG-----GATGAGCGCG 689
 QY 699 GGTGTGCGCAAGCTGGGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 758
 DB 690 CCGTGTGCGAGCGCTGTGAGACTTAAGGTGCGGAGCGCGCGCGCGCGCGCGCGCGCG 749
 QY 759 GCGCGCGCGCGCGCGCGCTTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818
 DB 750 GCTGTCCCGCTGCG 809
 QY 819 GCGCGCTACGAGCG 875
 DB 810 GCGCGAGCTGACG 869
 QY 876 GTCG 935
 DB 870 CCGTGAAG 929
 QY 936 CTCGAGCG 995
 DB 930 CACGAGCG 989
 QY 996 GCG 1055
 DB 990 GCGCGAGCG 1049
 QY 1056 CAACGAGCTGCG 1115
 DB 1050 CCGCGAGCTGCG 1109
 QY 1116 GCG 1175
 DB 1110 TCGCGTGCAGCG 1169
 QY 1176 GCGAGCTGCTCG 1235
 DB 1170 AGAGCTGACCTGAGTCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
 QY 1236 GCGAGCGCTGCG 1295
 DB 1230 GCGAGCGCTGCG 1289
 QY 1296 GCGTACCGCTGAGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1342
 DB 1290 CTCGACGATCATGATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1349
 QY 1343 -----TCAAGCG 1397
 DB 1350 AACCTCACCGAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1409
 QY 1398 CTCCTCGCGCATCTGAGTCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1457
 DB 1410 TGGGCG 1469
 QY 1458 GTCCCGCGCGTTCCTCCCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
 DB 1470 GTACCGCGTGAAGCG 1520
 QY 1518 CTTGATGCGCGAGTGAAGCAAGTCACTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1577
 DB 1521 CTATGAACGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1580
 QY 1578 CTGCG 1637
 DB 1581 CTGTGCGTGAAGCTGTGCG 1640

QY 1638 CAAGAGCTCATTAAGATCTACAG 1661
 DB 1641 CAAGAGCTCATTAAGATCTACAG 1664
 RESULT 8
 US-09-808-387-7
 ; Sequence 7, Application US/09808387
 ; Patent No. US2002013293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kalu Palm
 ; APPLICANT: Tonis Timmusk
 ; APPLICANT: Cernine Research
 ; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
 ; FILE REFERENCE: CEMRES.001A
 ; CURRENT APPLICATION NUMBER: US/09/808,387
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 1749
 ; TYPE: DNA
 ; ORGANISM: mouse
 US-09-808-387-7
 Query Match 23.4%; Score 391.6; DB 10; Length 1749;
 Best Local Similarity 55.9%; Pred. No. 1,1e-77;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;
 QY 99 GGTCTGCGCGAGCG 158
 DB 171 GCTCCCG 230
 QY 159 CCGCGCATCTGCG 218
 DB 231 CCG 290
 QY 219 GCG 278
 DB 231 CCG 350
 QY 279 GAGCG 338
 DB 351 GAGCG 410
 QY 339 CATCCCGAGTACCG 398
 DB 411 GCTGCCGAGTACCG 470
 QY 399 GCGCGAGAACCTGCG 458
 DB 471 GCTGAGGAGTTCG 530
 QY 459 GGTGTTCAGCGCGTGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
 DB 531 GGTCTTTCAGCGCGTGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
 QY 519 GCG 578
 DB 591 GAGCG 650
 QY 579 GAGCG 638
 DB 651 CAGCGAGCTGG-----TGTGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 701
 QY 639 CCGTCCCGCGCAGAGCAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
 DB 702 GCG 740
 QY 699 GGTGTGCGCAAGCTGGGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 758
 DB 741 CCGTGTGAGAGCTGTGCG 800

Oy	759	CGCGCGCGCGCGCGCATTTCCGTGGGAGGCCCGGATGAGGCGCCGCGGCGCGCGGTGTGGCC	818
Db	801	GCTTGGCGGCTGCGCCCGATTCGGCAGAACTCGTCAATTCTGACGACGCGCGCTGGCC	860
Oy	819	GGCGCTACTGAGAGCGCGACTGGCTTCGCTTCAAGCAACAGCGGAGCCCG---ACGTAGAGCT	875
Db	861	GGCGGAGCTGACCGCGCGACTGGCTTTCACAGCGCTTGGCGCGCGCGCGCGACGTCGCGAT	920
Oy	876	GTGGGCGGACCGCAAAAGTGGCTTGGCGACCGCGGCGCGGACGAGCTGGTCTT	935
Db	921	CCTGAGACGAGGACGAGTGGGCGCGCTTGAAGACAGGAGCGGACGAGCGCGCTGTCTT	980
Oy	936	CTCCGAGGCGCGCGCGCGCGCGCGCGGAGAGCTCTTGGTGAAGAGTGGCGCGTCCGCGGCT	995
Db	981	CACCAAGCGCGCGCTGTGAGCGTGGGCGGAGAACATCTTCAATCAAGGTCAAGCGCGCTCGGCGG	1040
Oy	996	GGCGGCGCGCGCGCGCGCGCTGGAGCTTCGAGATCAAGTCGTCGACACCGGCGCGTGTCAAGGCG	1055
Db	1041	GGGGGAGAGCGGCGCGCGCTGTCTTTCGGGGTCAACACGTCGTGACCTCTGGCAGCGTCGGCG	1100
Oy	1056	CAACGAGCTGCGCGCGCGGACCGACGCGCGCTGTGACCGCAAAAGTACTGGTGGTGGC	1115
Db	1101	CGCGGACCTGCGCTTCACGCGCGGAGCGCGCTGTGAGACCGCAAGGAGTTCTGGGCGGTGG	1160
Oy	1116	GGGCGCGGCGCGCGCGCGGAGCGCGCGCGGAGCGCGCTGACGTTCAACGCTCGCGCGCGCGG	1175
Db	1161	TGGCGTGGCGCGCGCGCTGTGACAGCGGAGCAATCTGGGCGCTGGTGTCAACGCGAGCG	1220
Oy	1176	CGAGGTGCTCTGGGCGATCAACGAGGCGTCGCGCGCGCGCGCGCTGTGTCGTCGACACAC	1235
Db	1221	AGAGGTGACCTGAGTCACAAAGGCGCGCGCGCGCGCGCATGACGTCGTGGTGAATGCTTC	1280
Oy	1236	GCAAGCGCTCTGGGCGCTTCTTGGCGCGCGCGCGCGCGCGTGGCGGCGACGTGCGTCTCT	1295
Db	1281	GCAAGCGCTCTGAGTGGCTCTTCAGCGCTGCATGCGCGCATCAAGCAGTCCGCACTTCG	1340
Oy	1286	CGGTACCTGAGTCGACCGCGCGGACCGACGACATCTCAACAGGTTCC-----	1342
Db	1341	CTCCACCATCAATGATGTGAAGGAGGTGGCCCATCTCTCCCTGCTCACTCGCTTCCATTC	1400
Oy	1343	-----TCAGCGGCTCCCGACGACGATGATGATTCAGATATGACCTTCAGTGTCAACGATTC	1397
Db	1401	AACCTCAACCGAGTGGCTTGGGCGATCCGCTGTGTGACCCCTTGGCTCAGACACTGGGGTTC	1460
Oy	1398	CTCCTCGGCACTGAGTCATCCCTGGTGAAGGCGCGCGCGCGCGCTCCCGCTAGCGCGCGGT	1457
Db	1461	TGGGCGCTTGAAGTGGCTGTGCTGAGGAGGACAGCGCCCAATTCACCTGTGAAGCTGCGCGGA	1520
Oy	1458	GTCCCGCGGTGTCTTCCCGACCGGAGCGCGGACGAGCATCAAGATGGCGAGTCAAGGTGG	1517
Db	1521	GTCAACCGGTGACCCCGAGGTCTGGGCGAG-----TGAAGTATGATGACCATTTG	1571
Oy	1518	CTTGCATGGGAGGTGAGACGAGTATCTACAGCTGTGGAACATGTGCTGTGGCAGAG	1577
Db	1572	CTATGAAACGCGAGTGAATCAGTATCTTACAGCTGTGGCCACATGTGCTGTGTATCTC	1631
Oy	1578	CTGCGGCTGGGCTCAAGCGACAGAGCCCGGCGCTGTCGCGCATCTGCGCGGCGCAT	1637
Db	1632	CTGTGGCTTGGCGCTCAAGAAAGGCGCTGCACGCGCTGCGCGCCCATCTGCGGCGCCAT	1691
Oy	1638	CAAGGACGTCAATTAAGATCTACAG	1661
Db	1692	CAAGGACATCAAGACCTTACCG	1715

RESULT 9
US-09-808-387-15
; Sequence 15, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk

```

1  APPLICANT: Cemines Research
2  TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
3  TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
4  FILE REFERENCE: CEMRES.001A
5  CURRENT APPLICATION NUMBER: US/09/808,387
6  CURRENT FILING DATE: 2001-03-14
7  NUMBER OF SEQ ID NOS: 48
8  SOFTWARE: FASTSEQ for Windows Version 4.0
9  SEQ ID NO 15
10 LENGTH: 1725
11 TYPE: DNA
12 ORGANISM: rat
13 US-09-808-387-15

```

Query Match	22.7%	Score 380.8;	DB 10;	Length 1725;
Best Local Similarity	54.7%	Pred. No. 2.7e-75;		
Matches 903; Conservative	0;	Mismatches 687;	Indels 60;	Gaps 5

OY	33	CCGAGGCCCA	CCGGGGGCGCTTC	TGGGCA	CCACCCGGCCGATGCTGGGCGCCCGGCGGAGG	92
Db	105	CCCGGTCCCTCT	CAACGATGCAT	CAAGAGAGAA	GGATTGGCCGCGCGCTTCAGG	164
OY	93	ACGCCGGGTCT	TGGGGGAGGGCG	CGCGCTTC	CAACGCGAGGCCAAAGGCATGTCGG	152
Db	165	TGGGGGGGCTCC	CGGGCAACACCGATGCT	CTTCCACCCCA	CTAAAGGGCTCCCAATCTCT	224
OY	153	GCTGAGCGGC	CACTCGCGCCGGGGCA	CACCGCGCAACAGCTTTC	GATATGGCTCACTT	212
Db	225	CATGACCTC	CAGCCACAAGAGCCGTC	CAAGAGCAAGCCAGCTTCTGTAA	TATGCACTACCTT	284
OY	213	CACCGAGCGGC	CAATCCGGCTGTAC	AGACAGATGGGCTGGCGCTGTATGGCGCTCGCC	272	
Db	285	CAGCAACGCGCC	CGCTCTCACTACAGAGCAAT	CAGGCTTGAAGATCACCAAGAGACATGG	344	
OY	273	TGGCTGAGAG	CGCGCGCTGGCTT	CGGCTTCA	CGCGCAACATCCGTCCTCATAGCGC	332
Db	345	CTGCTGAGAG	CGGGGCGCTGGCACTT	TGGCTTCA	CGCAGCAAGGACCTTCCGCATCACCC	404
OY	333	CCAGGACAT	CCCCAAGTACCGCTCCCGGCA	CTGGTCAACGGCGCGGCTTACTGGGCCAA	392	
Db	405	GAACTCACTG	CCCAAGTACCGCTCCCTGACCTGGTGTCC	CAAGATGCTTCTTGGGCCAA	464	
OY	393	GGCACTGGCC	CAGAACTGGCGCTTGC	CGACACGATGCTGGCTTACTGGGCCAACCGCA	452	
Db	465	AGCGTGGCTGA	GAGATTGTGC	CAACAGAGGCAACATATTCGCTCTTGGTGA	CAAGAA	524
OY	453	CGGCGCGTGT	TTTCAACAGCGTGAAC	GACCGGCGAGCGCGGTCTCTTTC	CACTGCGCGTGGC	512
Db	525	GGGCGGAGT	CTTTACCGGATTC	AAATGAACTCGGCTGC	CAATGCGTGTCTTTCAGCGGGTTCG	584
OY	513	CGTGGCGGCG	CGCGCTCTGGGGGCTCA	TTGATGTCTACCGGCAAT	CAACGAGAGTGCAGCT	572
Db	585	AAACGGCGGAC	CGCGCTCTGGGCGCTTGGTGA	CGTCTATATGACCTTCA	ACGAGGATGTTCAAGCT	644
OY	573	TCTGAGAG	CGCCTTCCGCTGA	CAACGCTGACCGCCCGCGGCTTC	CAGCCAGCCCGCTTCAG	632
Db	645	GCTAGACAGC	-----	GAAGCTGGTCTGCTTGA	CTGGCCGTAGAGCGCGCTCTTCAC	695
OY	633	CGGCTGCTG	CGCGCCAGACAGCC	CAAGACGGGCAACTTTC	CAACAAACAGAGCTGAGAA	692
Db	696	CGCGCTGGCG	CGCGCTGCTGCTG	CGGTGGAGGCGGAC	CAAGAGCGAGCC-----	742
OY	693	CAACGAGTGG	AGCAAGCTTGGGCA	CTTGGAGCGTGGAGCGCGCGCGGCCCAACCGCC	752	
Db	743	-----	TGCTGTGAGC	CTGTGTGCACTCA	ACCTGCTCAAGCAACAGCGCGCC	854
OY	753	AGCGCAGCG	CGCGCGCGCATTC	CGTGGCGGCGCCCGTGA	CGCGCCCGCGCGCGCTC	812
Db	795	CGACACGCG	CGCGCGCTGCG	CCGATCCCGCA	AACTCGCTCAAGCAACAGCGCGCC	854
OY	813	GTCGCGGCG	CTACTTGAAGCGCGAC	CTGTCCTTCAACGCAAC	ACGCGGGCCCG---AGCT	869
Db	855	GCTGCGGCG	CGAGCTCGACCGGCA	CTGCGCTTTCACAGCTCTG	CGCGCGCGCGCGCGCAAGT	914

```

Qy      870 GAGCTGTGCGCCGACCCGAAAGTGGCTTGGCGACACCGGCGCCGACGCGCCGCGACGCT 929
Db      915 CCGCATCTTGAACGAGCAGACGGTGGCGCGCTTAAGACAAGGCGCGACGAGCGCGCGCT 974
Qy      930 GATCTTCTCCGAGCGCCCGCTGCGCGCGCGCGAGAGCGCTTTCGTAAGAGTGGCGCGCTC 989
Db      975 GCTCTTCAACAGCGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1034
Qy      990 GGGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1049
Db      1035 AGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1094
Qy      1050 ACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1109
Db      1095 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1154
Qy      1110 GATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1169
Db      1155 GATGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1214
Qy      1170 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1229
Db      1215 GAGCGGAAAGCTGACCTTAATTCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1274
Qy      1230 CACCAAGCGAGCGCTTGGCGCGCTTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1289
Db      1275 GCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1334
Qy      1290 TCTCTCGGTACCGCTGAGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1346
Db      1335 CTTTGGCTTCAACATCATAGCTGAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1394
Qy      1347 -----CGGCTCCAGAGCGATAGTATTCAATATGACCTTCAAGTGA 1391
Db      1395 CACTTCAACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1454
Qy      1392 CCAATCTTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1451
Db      1455 CGATTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1514
Qy      1452 CCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1511
Db      1515 GCCGAGTACCGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1565
Qy      1512 GGTGTGCTTCAAGCGAGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1571
Db      1566 CATTTGCTATGAACGAGTGAATGATGATCATCATGATGATGATGATGATGATGATGATGATG 1625
Qy      1572 CCAAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1631
Db      1626 CTACTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1685
Qy      1632 GCCCATCAAGAGCTATTAGATTACAG 1661
Db      1686 CCCATCAAGAGCATCATCAAGAGCTACCG 1715

```

```

RESULT 10
US-09-808-387-5
; Sequence 5, Application US/09808387
; Patent No. US2002013293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemilnes Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PasteSeq for Windows Version 4.0

```

```

; SEQ ID NO 5
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-5

Query March 15.84; Score 264; DB 10; Length 1161;
Best Local Similarity 56.64; Pred. No. 2e-49;
Matches 567; Conservative 0; Mismatches 405; Indels 30; Gaps 3;

Qy      681 CAGACTGAGAAACAACAGGTGTGTGAGCGCAAGCTGAGCGCACTGAGCGCTGAGCGCGCGCC 740
Db      159 CAGAGCGAGACACCGCGCGCTCTGCGGTGAGCTATGAGACCTTCAACGTGTGCGCGCGCGCA 218
Qy      741 GGGGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
Db      219 CCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 278
Qy      801 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
Db      279 GCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 338
Qy      861 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
Db      339 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
Qy      918 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 977
Db      399 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 458
Qy      978 GGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1037
Db      459 GGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
Qy      1038 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1097
Db      519 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578
Qy      1098 AGAGTACTGAGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1157
Db      579 GGAATTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 638
Qy      1158 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217
Db      639 GGTGTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
Qy      1218 GCTGTGCGTGAACAACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1277
Db      699 GCTGTGCGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 758
Qy      1278 GGGCGAGCTGTGCTCTGAGTACCTGAGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1337
Db      759 GCAAGTCCGCAATCTCGCGCTCAGTATCTGCGCGAGGCGGAGTATCCGCTGATCTCCCTG 818
Qy      1338 GTCCCTCAGCGCGCT-----CCGAGAGCAATAGTATTAGATATGAC 1379
Db      819 CTCCCTGCTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 878
Qy      1380 CTTGAGTGTCAACAGTCTCTCTGCGCATGAGTCACTCTGAGTGAAGCGCGCGCGCGCGCGCT 1439
Db      879 GCTGAGACCGTGAAGCTTGTGCGCTCTGAGTGAAGCTTGTGAGTGAAGCGCGCGCGCGCG 938
Qy      1440 CCGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1499
Db      939 GCAAGTGAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 990
Qy      1500 TGGGAGTGAAGCGGTGTGCTTCAATGAGCGAGTGAAGCAAGTCACTTCAAGTGTGAACA 1559
Db      991 GATGAGTGAACCAATTTGCTATGAACAAGCGGTGAACAAGTGAACAAGTGAACAAGTGA 1049
Qy      1560 CATGTGCTGTGCGCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1619
Db      1050 CATGTGCTGTGCTAACCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109

```

QY 1620 CATCTCCGCGCGCCCATCAAGACCTCTTAAGATCTACAG 1661
DB 1110 CATCTCCGCGCGCCCATCAAGACCTCTTAAGATCTACAG 1151

RESULT 11

US-10-029-386-22715
; Sequence 22715, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22715
; LENGTH: 694
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121929.17
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: g114746644, EVALU 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P29503, EVALU 3.00e-10
; OTHER INFORMATION: EST_HUMAN HIT: AW340823.1, EVALU 0.00e+00
US-10-029-386-22715

Query Match 11.1%; Score 186.6; DB 13; Length 694;
Best Local Similarity 58.1%; Pred. No. 2.8e-32;
Matches 349; Conservative 0; Mismatches 249; Indels 3; Gaps 1;

QY 681 CAGCTCGAGAAACAACAGAGTGTGTGCAAGCTGGGCACTGGCGCTGGCGCGCC 740
DB 74 CAGAGGCGAGCAGACGGCGGCTCTCGGTGAGCTTATGCACTCAACGTGCGGGCGGGA 133
QY 741 GGGCCACCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
DB 134 CGCGCAGAGCG 193
QY 801 GGGCGCGCGGTGTGCG 860
DB 194 GCAACG 253
QY 861 GCCCG--ACGTGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
DB 254 CG 313
QY 918 CG 977
DB 314 CGAGCG 373
QY 978 GGTGGCGCGCTCG 1037
DB 374 GGTCAAGCG 433
QY 1038 CCCGCGCGGTCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1097
DB 434 CCCG 493
QY 1098 AAGATCTGGGT 1157
DB 494 GGAATTTCTGT 553
QY 1158 CAGCG 1217
DB 554 GGTGTCAACG 613

QY 1218 GCTGTGCTGTGACACACAGCAGCGCGCTGTGCGCTTCTTCCGCGCGCGCGCGCG 1277
DB 614 GCTGTGCTGTGACCGCTGTGCGCGCGCTTGTGATGTCTTCCGCGCGCGCGCGCG 673

RESULT 12

US-09-808-387-13
; Sequence 13, Application US/09808387
; Patent No. US2002012293A1
; GENERAL INFORMATION:
; APPLICANT: Kala Palm
; APPLICANT: Tonis Timusk
; APPLICANT: Cernine Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: TRANSRIPTION REGULATORS AND USBS THEREFOR
; CURRENT APPLICATION NUMBER: US/09/808,387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 876
; TYPE: DNA
; ORGANISM: mouse
US-09-808-387-13

Query Match 10.9%; Score 182.4; DB 10; Length 876;
Best Local Similarity 58.8%; Pred. No. 2.3e-31;
Matches 315; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 99 GGTCTGTGCGAGCG 158
DB 171 GCTTCCGCGCAGCG 230
QY 159 CGGCACTGTGCG 218
DB 231 CCTCAGCAGCAGCG 290
QY 219 GCGGCGCATCGCGCTGTACAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 278
DB 291 CGGCGCGGTCTCATCTACAGACAGTCAAGCTGAAGATCAACAGAGCAATGTGCTG 350
QY 279 GAGCG 338
DB 351 GAGCG 410
QY 339 CATCCCAAGTACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
DB 411 GCTGCGCAAGTACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
QY 399 GCCCGAAGACTGTGCG 458
DB 471 GCTTGAAGAGTTTTCAGAGCGAGCAATCATATGCTTCTTGTGAGTGAACAAGAGCG 530
QY 459 CCGTGTCTTACAGCGCGTGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
DB 531 GCTTCTTACAGCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 590
QY 519 CGGCG 578
DB 591 GGAACCGCGCTGTGCG 650
QY 579 GAGGCGCTTGTGTGACAGCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 634
DB 651 CAGCGAGCTGTGTGCG 706

RESULT 13


```

US-09-808-387-19
; Sequence 19, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemline Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 888
; TYPE: DNA
; ORGANISM: rat
US-09-808-387-19

```

Query Match 10.6%; Score 177; DB 10; Length 888;

Best Local Similarity 57.5%; Pred. No. 3.7e-30;

Matches 318; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```

Qy 33 CCCGAGCCACCGCGCGCTCTGCGCAACCGCGCGCTGCGGAGCCCGCGCCCGAGCG 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 CCCGATCCCTCTCACCAGTGCATCAGAGAGAGAGATTCGCCCGCGCTGTCAAG 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 93 ACGCCGCTCTGCGCGAGCGCGCGCTTCCAGCGCGAGCGCAAGAGCAAGAGCTGCG 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 TGGGGGCTCCCGCGCACACCGCTGCTTCCACCCCACTAAGGGCTCCCAAGTCT 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 153 GCTGACCGGCACTCGCGCGCGCGCACACCGCGCAACAGCTTGTGCAATGCGTCACTT 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 CATGACCTCAGCCAGAGCGCGCTCAGAGAGAGCGAGCTTGTATGCAATCACTT 284
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 213 CACGACGCGCCCATCGCGCTGTACAGAGAGTGGCGCTGCGCTGCGCGCGCGCG 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CAGCAACCGCCCGCTCATCTACAGCAAGTCAAGCTGAAAGTCAACCAAGAGCACTG 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 273 TGGCTGAGCGCGCGCGCTTGTGCGCTTCAACCGCGAGAGATCCGCTCATAGAGCG 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CTGCTGAGCGCGCGCGCTGCACTTGTGCTTCAACGAGAGACCTTCCGCTATCAACC 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 333 CCAGGACATCCCAAGTACGCTGCGCGCGAGCTGTCAAGCGCGCGCTACTAGGCGCA 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 CGACTCACTGCCCAAGTACGCTGCGCTGAGCTGTGTCCAGAGTGGCTTCTGGGCGCA 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 393 GGCATGCGCGAGAACTGCGCGCTGCGGAGACGCGTGTGCTTACTGCGCGCGCA 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 AGCGTGTCTGAGGAGTTTGCACAGAGGCGCAATCATTTGCTTCTGGTGTGACAGAA 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 453 CGGCGCGCTGTCTTACAGCGTGAAGAGAGCGCGCGTGTCTTCCACTCGCGGTGGC 512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 525 GGGCGGAGTCTTCTACCGGATCAAGTACGCGCTGCGCTGCTTCTTCAAGGGGTTCG 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 513 CGTGGGCGCGCGCTCTGCGCGCTCATTTATGTCTACAGGACATCAAGAGAGTGCAGCT 572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 585 AACGGCGGAGCCCGCTGTGGCGCTGTGTGAGAGTCTATAGGCTTCAACGCGGGTTCAGCT 644
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 573 TCTGAGAGCGCGC 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 645 GCTAGGAGACAGCC 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
US-09-808-387-17

; Sequence 17, Application US/09808387

; Patent No. US20020132293A1

; GENERAL INFORMATION:

; APPLICANT: Kaia Palm

; APPLICANT: Tonis Timmusk

; APPLICANT: Cemline Research

```

; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: rat
US-09-808-387-17

```

Query Match 10.6%; Score 176.8; DB 10; Length 1035;

Best Local Similarity 57.4%; Pred. No. 4e-30;

Matches 319; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

```

Qy 33 CCCGAGCCACCGCGCGCTCTGCGCAACCGCGCGCTGCGGAGCCCGCGCCCGAGCG 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 CCCGATCCCTCTCACCAGTGCATCAGAGAGAGAGATTCGCCCGCGCTGTCAAG 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 93 ACGCCGCTCTGCGCGAGCGCGCGCTTCCAGCGCGAGCGCAAGAGCAAGAGCTGCG 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 TGGGGGCTCCCGCGCACACCGCTGCTTCCACCCCACTAAGGGCTCCCAAGTCT 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 153 GCTGACCGGCACTCGCGCGCGCGCACACCGCGCAACAGCTTGTGCAATGCGTCACTT 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 CATGACCTCAGCCAGAGCGCGCTCAGAGAGAGAGAGCTTGTATGCAATCACTT 284
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 213 CACGACGCGCCCATCGCGCTGTACAGAGAGTGGCGCTGCGCTGCGCGCGCGCG 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CAGCAACCGCCCGCTCATCTACAGCAAGTCAAGCTGAAAGTCAACCAAGAGCACTG 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 273 TGGCTGAGCGCGCGCGCTTGTGCGCTTCAACCGCGAGAGATCCGCTCATAGAGCG 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CTGCTGAGCGCGCGCGCTGCACTTGTGCTTCAACGAGAGACCTTCCGCTATCAACC 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 333 CCAGGACATCCCAAGTACGCTGCGCGCGAGCTGTCAAGCGCGCGCTACTAGGCGCA 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 CGACTCACTGCCCAAGTACGCTGCGCTGAGCTGTGTCCAGAGTGGCTTCTGGGCGCA 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 393 GGCATGCGCGAGAACTGCGCGCTGCGGAGACGCGTGTGCTTACTGCGCGCGCA 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 AGCGTGTCTGAGGAGTTTGCACAGAGGCGCAATCATTTGCTTCTGGTGTGACAGAA 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 453 CGGCGCGCTGTCTTACAGCGTGAAGAGAGCGCGCGTGTCTTCCACTCGCGGTGGC 512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 525 GGGCGGAGTCTTCTACCGGATCAAGTACGCGCTGCGCTGCTTCTTCAAGGGGTTCG 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 513 CGTGGGCGCGCGCTCTGCGCGCTCATTTATGTCTACAGGACATCAAGAGAGTGCAGCT 572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 585 AACGGCGGAGCCCGCTGTGGCGCTGTGTGAGAGTCTATAGGCTTCAACGCGGGTTCAGCT 644
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 573 TCTGAGAGCGCGCTTC 588
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 645 GCTAGGCTTCAACATC 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

US-09-808-387-11

; Sequence 11, Application US/09808387

; Patent No. US20020132293A1

; GENERAL INFORMATION:

; APPLICANT: Kaia Palm

; APPLICANT: Tonis Timmusk

; APPLICANT: Cemline Research

; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF

; FILE REFERENCE: CEMRES.001A

; CURRENT APPLICATION NUMBER: US/09/808.387

; CURRENT FILING DATE: 2001-03-14

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: mouse
US-09-808-387-11

Query Match 10.5%; Score 176.4; DB 10; Length 1035;
Best Local Similarity 60.0%; Pred. No. 4.9e-30;
Matches 294; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```
QY 99 GGTCTGGGCGAGGCGCGCGCTTCAACGGCGCAGGCCAAAGCAGAAAGCTGGCGCTGGA 158
   |||||
Db 171 GCTCCCGGCGACGCCGCTGCTTCCACCCCACTAAAGGCTCCAGATCCTCATGGA 230
   |||||
QY 159 CGGCCACTGCGCGCGCGGCCACAGCGGCAACAGCTTGTGCAATGGGCTCAAGTTCAAGCA 218
   |||||
Db 231 CTTAGCCACAAAGGCGGCTCAAGAGGCAAGCCAGCTTGTCAATGTCATCACTTCACTAA 290
   |||||
QY 219 GCGGCCCATCCGCGCTGTACAGAGAGGTGGCGCTGCGCTGTGTGGCCGTGCGCCCTGGCTG 278
   |||||
Db 291 CCGCCCGGTGTCTATCTACAGAGAGTCAAGGCTGAAGATCAACAAGAGCAATGCTGTG 350
   |||||
QY 279 GAGCGGCGCGCTGCGCTTGCACCGCGCAGATCCGTGCTCATGAGGCCCAAGGA 338
   |||||
Db 351 GAGCGGGGCGCTGCGCTTGCACCGCAAGAGACCTTCCGCAATCCACCCGACTC 410
   |||||
QY 339 CATCCCAAGTACGCGTGCAGGAGCTGTGTCAGCGGCGCGGCTACTGGGCCAAAGGCACT 398
   |||||
Db 411 GCTGCCAAGTAGCGCTGCTGACCTGTGTCTCAAGAGTGTCTTGTGGCCAAAGCAATT 470
   |||||
QY 399 GCCCGAGAACTGGCGCTGCGCGACAGGTGTGCTGTGCTTACTGGGCCGACCGCCACGGCCG 458
   |||||
Db 471 GCTGAGAGAGTTGCCAGAGGCGCAACATATTGCTTGTGGTGGACAGAGAGGCGCG 530
   |||||
QY 459 CGTGTCTACAGCGTGAACAGACGGCAGCGGTCTTTCACCTGCGCGGTGGCGGTGG 518
   |||||
Db 531 CGTCTTCTACCGGATCAATGAGTCAAGCTGTATGCTTTTCTTCAGTGGGGTCCGAGACGGT 590
   |||||
QY 519 CGGCCGCGCTGGGCGCTCATTTGATGTCTACGGGATCACCGACGAGGTGCAAGTCTTGA 578
   |||||
Db 591 GGAACCGCTGTGGGCGCTGTGAGAGTCTACGGGCTCACGGGGGTGTCAAGCTGTAGG 650
   |||||
QY 579 GAGCGGCTTC 588
   |||||
Db 651 CTCACCAATC 660
   |||||
```

Search completed: January 20, 2004, 20:10:06
Job time : 586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2004, 15:07:52 ; Search time 3269 Seconds

(Without alignments)
12453.349 Million cell updates/sec

Title: US-09-808-387-21

Perfect score: 1675
Sequence: 1 atgggcaacacgctgcacg.....ctacagccatagcctagcc 1675

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*
1: em_earth:*
2: em_earth:*
3: em_earth:*
4: em_earth:*
5: em_earth:*
6: em_earth:*
7: em_earth:*
8: em_earth:*
9: gb_earth:*
10: gb_earth:*
11: gb_earth:*
12: gb_earth:*
13: gb_earth:*
14: gb_earth:*
15: em_earth:*
16: em_earth:*
17: em_earth:*
18: em_earth:*
19: em_earth:*
20: em_earth:*
21: em_earth:*
22: em_earth:*
23: em_earth:*
24: em_earth:*
25: em_earth:*
26: em_earth:*
27: em_earth:*
28: gb_earth:*
29: gb_earth:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	30.9	573	12	B1848332
2	509.6	30.4	517	9	A1571301
3	492.8	29.4	644	9	AA850748
4	362.8	21.7	470	9	AA049442

Result No.	Score	Query Match	Length	DB ID	Description
5	343.8	20.5	464	9	A1113327
6	339	20.2	3212	11	AK078529
7	296.8	17.7	703	3	B0238931
8	288.6	17.2	377	9	A1579407
9	278.6	16.6	593	10	BB097995
10	275.4	16.4	392	4	AM529983
11	264.4	15.8	665	12	BM440106
12	263	15.7	304	12	B1404107
13	261.8	15.6	472	9	AA851139
14	242.2	14.5	311	9	A1577954
15	220.2	13.1	362	28	A2953956
16	210.6	12.6	355	28	A2341570
17	206	12.3	537	12	B134290
18	199.2	11.9	527	10	BB649566
19	194.2	11.6	255	9	AA957126
20	194.2	11.6	318	9	AA489411
21	185.8	11.1	623	14	CA941769
22	185.8	11.1	626	12	B1712399
23	185.8	11.1	681	13	BX098217
24	185.8	11.1	715	14	U69206
25	182.6	10.9	578	10	BB656496
26	182.4	10.9	566	13	BU952430
27	182.4	10.9	588	13	BQ269642
28	181	10.8	570	14	CA948476
29	180	10.7	568	14	CA941491
30	178.4	10.7	984	13	BU452591
31	177.6	10.6	615	12	B1712632
32	175.8	10.5	589	14	CA771909
33	175.8	10.5	631	14	CA774789
34	175.4	10.5	2093	11	AK044464
35	172.2	10.3	661	10	BF147610
36	171.8	10.3	1060	11	AK010787
37	171.8	10.3	4034	11	AK046382
38	169.4	10.1	573	13	CA843851
39	169.4	10.1	645	13	BQ269642
40	167.4	10.0	431	12	BM724707
41	167.4	10.0	431	12	BM930417
42	167.4	10.0	447	10	BP591302
43	166.8	10.0	436	9	A1220894
44	163.2	9.7	559	14	CA941339
45	162.8	9.7	215	10	BB863961

ALIGNMENTS

RESULT 1
LOCUS B1848332
DEFINITION 470743 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1848332
VERSION B1848332.1 GI:15960851
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 573)
Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Cassas,B., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Pettee,G., Holt,I., Karamycheva,S., Liang,P., Quackenbush,J. and
Keefe,J.W.

TITLE JOURNAL MEDLINE PUBMED
COMMENT Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

CONTACT: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366

Df	277	GGCCGCGTGTCTACAGAGTGAAAGAACGCGGCAGCCGGTGTCTTCATCTGCGGGGTGGCC	218
Oy	514	GTGGGCGGCCCGCCTCTCTGGGCGCCTCATTTATGTCTTAAGCATCAACGAAGTGTCACTT	573
Df	217	GTTGGCGCGCCCGCTCTGTGGCGCCTCATTTATGTCTTAAGCATCAACGAAGTGTCACTT	158
Oy	574	CTGAGAAGCGCCTTCTGCTGACAGCGTGAAGCGCCGCGCGCCTCAGCAGAGCCCGCTTACG	633
Df	157	CTGAGAAGCGCCTTCTGCTGACAGCGTGAAGCGCCGCGCGCCTCAGCAGAGCCCGCTTACG	98
Oy	634	GCTTGCTCTCCGCCAGCAGCAGCACGCGGCCAATTTCACAACAACGACTCGAAGAC	693
Df	97	GCTTGCTCTCCGCCAGCAGCAGCACGCGGCCAATTTCACAACAACGACTCGAAGAC	38
Oy	694	AACCAAGTGTGTGCCCAAGCTGTGGCCCACTGTGCGCTGG	730
Df	37	AACCAAGTGTGTGCCCAAGCTGTGGCCCACTGTGCGCTGG	1
RESULT 3			
LOCUS	AA850748	644 bp	mRNA linear EST 30-APR-1998
DEFINITION	EST193516 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone ROVAK34 3' end, mRNA sequence.		
ACCESSION	AA850748		
VERSION	AA850748.1	GI:2938288	
KEYWORDS	EST.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 644) Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., AUTHORS Kerlavages,A.R. and Adams,M.D. Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat TITLE Gene Index Unpublished COMMENT Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel.: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@igr.org Seq primer: M13-21.		
JOURNAL			
FEATURES			
source	location/Qualifiers		
	1..644		
	/organism="Rattus sp."		
	/mol_type="mRNA"		
	/db_xref="ATCC (Inhost):2009149"		
	/db_xref="ATCC (Inhost):2009149"		
	/db_xref="taxon:10118"		
	/clone_id="ROVAK34"		
	/cdoone_jib="Normalized rat ovary, Bento Soares"		
	/note="Organ: ovary; Vector: pRTT3Bac; Site_1: EcoRI; Site_2: NotI"		
BASE COUNT	87 a 218 c 226 g 113 t		
ORIGIN			
Query Match	29 4%;	Score 492.8;	DB 9; Length 644;
Best Local Similarity	85.6%;	Pred.No.1.5e-75;	
Matches 548; Conservative	0;	Mismatches 92;	Indels 0; Gaps
Df	953	GAGCCGAGAGAGCCTCTTCTGTAAGAGTGGAGCGCCGTCCGAGGCTGAGCGCCCGAGCGGC	1012
Oy	5	GCCCCGGGGAAGCCTGTGCTGCTGGAAGTGGAGCGCCCGAGGCTGAGCGCCCGAGCGGC	64
Df	1013	TGGCTTGGGATCACTGTGTGCGAACCGGAGCGTGTACAGGCCCAAAGAGTGTGCGCCGCG	1072
Oy	65	TGGCTTGGGATCACTGTGTGCGAACCGGAGCGTGTACAGGCCCAAAGAGTGTGCGCCGCG	124
Df	1073	ACCAGAGCGCGTGTCTCAACCGCAAGAGTACTGTGTGTGTGGCGGCGCGCGCGCGCGTGC	1133
Oy	125	ATCCCGCTGCGTGTCTGAGCCCGCAAGATATCTGAGTGTGTGCGCGCGCGCGCGCGTGC	184

OY		1133	CGAGCCGCGCGCAGACGGCTCAGCTTCAACGTGTGGCGCGCGCGCGCGACATGCTCTCTGGGCA	1192
Dd		185	CGAGCGGAGCGCAGCGACTGAGCTTTCAACGCTTGCGAACCGCGCGCGAATCTCTGCGCGG	244
OY		1193	TCAAACGCGCCTCCGCGCGCGCGCTGCTGTGCCTGACACAACGACGACGCGCTTGCGCT	1252
Dd		245	TGAACCGGAGCGCGCGCGGAGACGCTTGCTGTGGTAGACCTCGACAGGCGCTCTGGGCT	304
OY		1253	TCTTCGCCCTGCGCGCGCGCGCTGCGCGGCGCACGCTGCGTCTCTCCGGTAACCTTGACATCCA	1312
Dd		305	TCTTCGCTGTGCGCGCGGCTGTGTGGCGGCTAGCTGTGCTCTCGGGGACCGTACAGTCG	364
OY		1313	GCCCTGTGCACACGACTCATCAAGGGGCTCTCAAGCGGCTCCAGAGAAATAGATTGAG	1372
Dd		365	GTCTTAGGCGCACACTCCATCAAGGGTCTTCAATGGCTCTCAGAATGACAGCATTCGG	424
OY		1373	ATAATGACCTTCACTGTTCMAACAGTCTCTCTCGGCATGTAGTCAATCTCTGTGTACGCGCC	1432
Dd		425	ACATGACCTTCGGGGGTCAACCAATGCGTATCACATGACATGAAATGCTCTCTGTGTACAGCCC	484
OY		1433	CGAGCTCCCGCTGAGCGCCCCCGCGGTGCCCCCGTGTTCGCCACCGGAGCCGCGAGGCA	1492
Dd		485	CGAGCTCCCACTGAGTCCCCCAGATGCCCCGCGCTTCTGTGACCAGAACCGGCGCGCA	544
OY		1493	TCAAAGATGCGCAGTGCACGCTGTGCTTTCAGATGGCGAGGTGACACCGCTCATCTACAGT	1552
Dd		545	GCAGGAATGAGAGATGACCGGTGTCTTTCGACAGCAGAGGTGACACCGCTCATCTACAGT	604
OY		1553	GTGACACATATGTGCTGTGCGACAGCTGCGGCGCTTGCGCT	1592
Dd		605	GTGACACATATGTGCTGTGCGACAGCTGCGCGCTTGCGCT	644
RESULT 4				
AM049442				
LOCUS				
DEFINITION				
UI-M-BH1-ams-b-10-0-UI.s1 NIH BMAP M 52 Mus musculus cDNA clone				
AM049442				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Mus musculus (house mouse)				
Bukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.				
1 (bases 1 to 470)				
Bonaldo,M.F., Lennon,G. and Soares,M.B.				
Normalization and subtraction: two approaches to facilitate gene				
discovery				
Genome Res. 6 (9), 791-806 (1996)				
JOURNAL MEDLINE PUBMED COMMENT				
97044477				
8889548				
Contact: Chin, H				
National Institute of Mental Health				
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD				
20892-9643, USA				
Tel: 301 443 1706				
Fax: 301 443 9890				
Email: msthemali.nih.gov				
Oligo-dt track not found, Not I site shown in beginning of sequence				
is likely internal to the message. cDNA Library Preparation: M.B.				
Soares Lab Clone distribution: NIH BMAP cDNA clones will be made				
determines the means that is soon to be determined. When NIH				
record will be updated accordingly when that means is determined.				
Seq primer: M13 Forward				
POLYA=No.				
FEATURES				
source				
location/Qualifiers				
1..470				
/organism="Mus musculus"				
/mol_type="mRNA"				
/strain="C57BL/6J"				

/db_xref="taxon:10090"
 /clone="UI-M-BH1-ams-b-10-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M S2"
 /note="Vector: pTZ19-Lac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP M S2 library is a subtracted library derived from
 NIH BMAP M S1, which in turn is a subtracted library
 derived from a mixture of normalized libraries from ten
 regions of the mouse brain (cerebellum, brain stem,
 olfactory bulb, hypothalamus, cortex, amygdala, basal
 ganglia, pineal gland, striatum, hippocampus). The driver
 used for subtraction consisted of a pool of 5,000 clones
 from the NIH BMAP M S1 library and a pool of 2,000 clones
 obtained from non-normalized and normalized mouse brain
 spinal cord libraries.
 TAG LIB=NIH BMAP M S2
 TAG TISSUE=hypothalamus
 TAG SEQ=CCGTA"

BASE COUNT 62 a 162 c 168 g 78 t
 ORIGIN

Query Match 21.7%; Score 362.8; DB 9; Length 470;
 Best Local Similarity 85.7%; Pred. No. 4e-53;
 Matches 403; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 943 CGCCGCGCTGCGCCGCGGCGAGAGCTCTTGTGAGAGTGGCGGCTCGGGGCTGGCGGCG 1002
 DB 1 CGGCGCGCTGCGCGGCGGAGAGCTGTGAGAGTGGCGGCGCGCGGCGCTGGCGGCGA 60
 QY 1003 CGCGGCGCGCTGCGCGCTTGGGATCACTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1062
 DB 61 CGCGGCGCTGTGCGCTTGGGATCACTGTGAGACCGCGGCGGTCTACGCGCAACGAG 120
 QY 1063 CTGCGCGCGGCGAGCCGAGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1122
 DB 121 CTGCGCGCGGAGTCCGCTGCGGATCGGACCGGAGATATCGGAGTGGCTCGCGCC 180
 QY 1123 GGGCGCGTCCGAGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1182
 DB 181 GGGCGCGTCCGAGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 240
 QY 1183 CTGCGCGGATCAACGCGGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1242
 DB 241 CTGCGAGCGGTGAACGCGGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 300
 QY 1243 CTGCGGCGCTTCTTCCGCGGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1302
 DB 301 CTGCGGCGCTTCTTCCGCGGCGGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 360
 QY 1303 CTGCGAGTCCAGCGGCGGCGGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1362
 DB 361 CTACAGCTCCAGTCTTCAAGACGATGACCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 420
 QY 1363 AGGATTCAGATATGACCTTCACTGATCAACGAGTCCCTCGGATCTGA 1412
 DB 421 AGCGATTCAAGACATGACTTGGCGGTCAACGAGTCTCTCAGACATCAGA 470

RESULT 5
 A1113327 464 bp mRNA linear EST 11-FEB-1999
 LOCUS A1113327
 DEFINITION UI-R-C2p-nr-f-12-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
 ACCESSION A1113327
 VERSION A1113327.1 GI:3513276
 KEYWORDS EST
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 464)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBR, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA library Preparation: M.
 Patricia Ronaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics The following repetitive elements were
 found in this cDNA sequence: 39-65, >GC_rich_low_complexity
 Seq primer: M13 Forward.

FEATURES
 source
 Location/Qualifiers
 1..464
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-nr-f-12-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C2p"
 /note="Vector: pTZ19-Lac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dt track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' BstBI had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Ronaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)."

BASE COUNT 54 a 157 c 173 g 80 t
 ORIGIN

Query Match 20.5%; Score 343.8; DB 9; Length 464;
 Best Local Similarity 85.1%; Pred. No. 7.5e-50;
 Matches 384; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 943 CGCCGCGTGGCGCCGCGGCGAGAGCTCTTGTGAGAGTGGCGGCTCGGGGCTGGCGGCG 1002
 DB 1 CGGCGCGCTGCGCGGCGGAGAGCTGTGAGAGTGGCGGCGCGCGGCGCTGGCGGCG 60
 QY 1003 CGCGGCGCGCTGCGCGCTTGGGATCACTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1062
 DB 61 CGCGGAGCTGTGCGCTTGGGATCACTGTGAGACCGCGGCGGTCTACGCGCAACGAG 120
 QY 1063 CTGCGCGCGGCGAGCCGAGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1122
 DB 121 CTGCGCGCGGAGTCCGCTGCGGATCGGACCGGAGATATCGGAGTGGCTCGCGCC 180
 QY 1123 GGGCGCGTCCGAGCGGCGGCGGCGGCTGTGAGACCGCGGCGGTCTACGCGCAACGAG 1182

Db 181 GGGCCCTGGCCAGCGAGCGAGCACTGAGCTTCAAGCTGCGACCGGGGCGAGCTC 240
 Oy 1183 CTCCTGGGACATCAACGGGCGTCCGCGCGCGCTGTGTGCTGACACACCGACGCG 1242
 Db 241 CTGCTGGCGGTTGAACGGGCGCGCGCGCGGAGCGCTTCTGTGCTGACACCTCCGACGCG 300
 Oy 1243 CTCCTGGGCGCTTCTTTCGCGGTCGCGCGCGCGCTGCGCGGCGACGCTGCTCTCGGTACC 1302
 Db 301 CTCCTGGGCGCTTCTTCTGCTGTGCGCGGTGTGTGCGGCGGTACGCTGCTCTCGGCGAC 360
 Oy 1303 CTCGAGTTCAGCGCTTGGGACACACGATCTGACGAGTCTCCGACGCGCTCCACAGCAT 1362
 Db 361 GTACAGTTCGCGCTTGGAGCGCACCACTGACGAGTCTTCAAGTGTCTTCAGAGTAC 420
 Oy 1363 AGTGAATCAGATGATGACCTTCAGTGTCAACC 1393
 Db 421 AGCGATTCGAGCATGACCTTGGGCGTCAACC 451

RESULT 6
 AK078529
 LOCUS
 DEFINITION Mus musculus 12 days embryo female mullerian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:682043L06 product:hypothetical protein, full insert sequence.
 AK078529
 AK078529 1 GI:26097916
 HTG; CAP trapper.
 Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Mech. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitenai, T., Tashiro, H., Itoh, M., Suni, I., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Fletcher, C., Fujita, M., Gariboldi, M., Guerinich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaeys, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kontouki, S. and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851

REFERENCE
 AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 3212)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hamagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koude, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshida Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-research.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: <http://genome.gsc.riken.go.jp/>
 URL: <http://fantom.gsc.riken.go.jp/>
 FEATURES
 SOURCE
 1..3212
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM:DB:682043L06"
 /db_xref="taxon:10090"
 /clone="682043L06"
 /sex="female"
 /tissue_type="mullerian duct includes surrounding region"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 1..3212
 /note="hypothetical protein (evidence: decoder, longest-ORF)"
 ORIGIN
 BASE COUNT 562 a 993 c 943 g 714 t
 Query Match 20.2%; Score 339; DB 11; Length 3212;
 Best Local Similarity 56.3%; Pred. No. 5.6e-49;
 Matches 794; Conservative 0; Mismatches 555; Indels 62; Gaps 6;
 Oy 275 GCTGAGCGGCGCGCTGCTGCTTCAACCGGCGACGATCCGCTGCTATGAGCGCC 334
 Db 51 GCTGAGCGGCGCGCTGCTGCTTCAACCGGCGACGATCCGCTGCTATGAGCGCC 110

OY	335	AGGACATCCCCAAGTAAGCCCGGACCCCGACCCGGTCAAGCCGGCCGAGCTTACGGGCGCAAGG	394
Db	111	ACTGCGTGGCCCAAGTAAGCCCGTGGCCCTGACCTGGGTCTCAAGATGGGCTTCTGGGCGCAAGG	170
OY	395	CATGCGCCGAGAACTTGCGCGCTGCGCGCAACAGGTGCTGGACTTACTTGAGCCGACCCGACCG	454
Db	171	CATTGGCTGAGGAATTTTGGCCAAACAGAGGCAACATCATTTGGCTCTTGGGGTGAGCAAGAAAG	230
OY	455	GCCGCGTGTTCACAGCGTGAACGACGCGAGCCGGGTGCTTTCACCTGCGGCGTGGCCG	514
Db	231	GCCCGTCTTCTACCGGATCAATGAGTCAAGCTGGCTAATGCTTTTCTTCAGTGGGGTCCGGA	290
OY	515	TGGGCGGCGCCGCTCTGGGCGCTCAATGATGTCTAAGGATATACCGACGAGGTGAGCTTC	574
Db	291	CGGTGGAACCGGCTCTGGGCGCTTGAGACGTCTACGGGCTTACCGCGGGGGTGTCCAGCTGC	350
OY	575	TGAGAGAGCGCTTGCGCTGACACGCTGACGCGCCGCGCGCTCAGCGCAAGGCGCGCTTCAGCG	634
Db	351	TGAGACAGC-----GAGCTGGGTCTGCGCCGACTGCGGCGCGGCTCTTCACCG	401
OY	635	CTTGCCTTGCSCCGACGACCGACGACCGCGGCGCAACTTCGACCAACAGAGCTCGAAGACA	694
Db	402	CGTTCGCGCGCGCGCTGCGCTGCGGTGGTCGAGGCGGAT-----GAGGCGC	443
OY	695	AACAGAGTGTGTGGCCAAAGTGGGCGCACTGGGCGCTGGGCGCGGCGCGCGCGCGCGCGCG	754
Db	444	GCTGTGCTGTAGGCTGTGGCCACTGACCTCAACGTGCGGGAGCGG--ACGGGACGACGGGCT	501
OY	755	CGACAGCGCGGCGCGCGCGCATTCCTGATCGGAGCCCGCTGAGACGCGCGCGCGCGCGCGT	814
Db	502	CACCGCGCTGCGGCTGCGCGCTGATCCCGGCAAACTCGCTCAATTCCTCAGACAGCGCGCGC	561
OY	815	CGCGGCGCTACTGAGAGGCGGACCTGCGCTT-----CCAGCGAACACGCGGCGCCGACG	868
Db	562	TGCCGCGCGACACTGACGCGGACCTGCGCTTTCACAGCGCGGCTGCGCGCGCGCGCGACGT	621
OY	869	TGAGCGCTGTGGCGCGACCGGCAAAATGGGCTGTGACACCGCGGCGCGACCGCGCGCGCGC	928
Db	622	TCCGCATGTGTGAGACGAGCAGACGCGGTGGCGCGCTTGAGAACAGGCGCGCACAGCGCGCGC	681
OY	929	TGGTCTTCTCGAGAGCGCCGCGTGGCGGCGCGCGGAGAGCTCTTCGTGAGGTGGGCGGCTC	988
Db	682	TGCTCTTACAGCGCGGCTGTGGCGCGTGGCGGAGACATCTTATCAAGGTCAAGCGGCT	741
OY	989	CGGCGCTTGGCGGCGCGCGCGCGCGCGCTGCGCTTCGAGATCAAGTGTGCGACCGGGCGTGC	1048
Db	742	CGGGCGGGGGGAGAGCGGGCGGCGGTCTTCGGGGTCAACACGATGTACCCCTGGACGC	801
OY	1049	TACGGCCCAAGAGCTGCGCGCGCGCGCACGACGCGTCTGATCGCAAAAGATCTGAGG	1108
Db	802	TGCGGCGCGCGACCTGCGCTTCACGCCCCGAGGCGCTGTGTGACCGCAAGAGTTCCTGGG	861
OY	1109	TGGTGGCGCGGCGCGGCGCGCGCGCGGAGCGGGGAGAGCGGTGAGCTTCACGCTGCGGC	1168
Db	862	CGGTGTGTGGGTGGCGCGCGCGCTCTGCAAGAGGAGATCTTGGGCGTGGGTCAAGC	921
OY	1169	CGCGCGGAGAGTGTCTCTCTGGGCAATCAACGAGCGTCCGCGCGCGCGCGCTGTGGTGGTGC	1228
Db	922	CGAGAGGAGAGTGAACCTGAGTCAACAGCGGCGCGCGCGCGCGAGTGAAGGTGTGGTGG	981
OY	1229	ACACCAAGAGCGGCTCTGGGCGCTTCTTCGCGCGTGGCGGCGGCGGTGCGGCGGACGCTGC	1288
Db	982	ATGCTCTCCAGACCCCTCTGGATATGCTTTCAGCTTGCATAGGCGGCGCATCAACGAGGTCCG	1041
OY	1289	GTCCTCTGGATACCTGTGAGTCCAGGCGCGCGGACCAAGATCCATCAAGGGT-----	1339
Db	1042	TCTCTGGCTCCACATCATGATGTAACGGGGTGGCCATCTCTCCCTGCTCAGCTGCT	1101
OY	1340	-----CCTTCAGCGGCTCCGAGACGATAGTGAATTCAGATTAAGATTCCTTCACTGTCA	1390
Db	1102	CGACTCCAACTCAACCGAGTGGCGTGGGATACCGGCTGTGAGACCCCGTGTGACAGCACT	1161
OY	1391	ACCACTCTCTCGGACATCTGAGTATCTCCGTGTGACGAGCCCGGACCTCCCGCTGAGCC	1456

Db	1162	GCGGATCTGGAGCCCCCTTAGTGGCTCTGTCTGAGAGGAGCAGAGCCCCCACTCACTCTGTAGGCC	1221
Qy	1454	CCCCGGGTGTCCCCCGGTATTCTCCCAACGGAGCGCGAGGCATCAAGAAATGGGAGAGCA	1510
Db	1222	TGCCCAAGTCAACCGGTGACCCCAAGTCTG-----GGCCAATGAGATGATGAATGCA	1272
Qy	1511	CGGTGTCTTCGATTGGCGAGGTGGACAAGGTCAATCTTACAAGTGTGAAACAATGTGCTGT	1570
Db	1273	CCATTTCCTATTGAACAAGCAGGTGACAACATCTCATCTTACAGTGTGGCCACATGTGCTGT	1332
Qy	1571	GCACAAGCTCGGCGCTCGGCTCAAGCAAGGCCCGGCGCTGTGCCCCCATCTGCCGCG	1630
Db	1333	GCTACTCTGTGTGGCTTGGCGCTCAAGAAAGGCCCTGCAAGCTGTGCCCCCATCTGCCGCTC	1392
Qy	1631	GGCCCATCAAGAGACGTCAATTAAGTCTACAG	1661
Db	1393	GCCCCATCAAGACATCATCAAGACCTTACCG	1423

RESULT 7	BU238931	703 bp	mRNA	linear	EST 26-NOV-2002
LOCUS	BU238931				
DEFINITION	BU238931 603322904.F1 CSBECN33	Gallus gallus	CDNA	clone CH872512.5	mRNA
ACCESSION	BU238931				
VERSION	BU238931.1	GI:25484991			
KEYWORDS	BST.				
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
1 (bases 1 to 703)	Boardman, P.R., Sant-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Ford, M.T., Rickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.	A Comprehensive Collection of Chicken CDNA	Curr. Biol.	12 (22)	1965-1969 (2002)	
			22335534			
			12445392			Contact: Simon Hubbard

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1BD, UK
Tel: 01612008930
Fax: 01612360405
Email: Simon.Hubbard@umist.ac.uk
Location/Qualifiers
1..703

```

/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton line 151"
/db_xref="taxon:9031"
/clone="CNS175192"
/bex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="CBROCMN33"
/note="Organ: liver; Vector: Bluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dt) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the Bluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

```

BASE COUNT 150 a 228 c 202 g 122 t 1 others
ORIGIN

Query Match 17.7%; Score 296.8; DB 13; Length 703;
Best Local Similarity 72.8%; Pred. No. 9.9e-42;
Matches 456; Conservative 0; Mismatches 152; Indels 18; Gaps 5;

QY 1 ATGGGCAACACGGTGCACCGGACCTTGCAGACCCGAGCCACCGGCGGCTCTCTGACC 60
DB 77 ATGGGCAACACGGTGCACCGGACCTTGCAGACCCGAGCCACCGGCGGCTCTCTGACC 136
QY 61 ACCGGGCGGCTGTGGGGCC-----CGGCCCCGAGCCAGCCGCGGCTCTCTGAGC--- 108
DB 137 AGCCGCTCCCTACTAGACGCTTCCCAACGACGACCTAGGCGACGCTCAGTGTCTCTGTCTC 196
QY 109 ---GAGGCGCGCGCTTCCACGCGGAGGCGCAAGAGGCAAGAGCGTGGCTGGACGGCCAC 165
DB 197 ACGAGGCGCGCTTCCCTTCCACCTCCACGCGGAGGCGCAAGAGCGTGGCTGGACCGAC 256
QY 166 TCGCGCGGCGGCGCACGCGGCGCAACAGCTTCTGCAATGGCGTCACTTCAACGAGGCGCC 225
DB 257 TCCCGACGGGCGCACGAGGAGGAGCAAGCTTCTGTATCGGATCACCTTCAACGAGGCGCC 316
QY 226 ATCCGCGCTGTACGAGGAGGCTGCGCTGCGCTGTGGCGCGCTGGCGCGCTGGAGCGGC 285
DB 317 ATCCAGCTCTACGAGAGAGGTGCGGCTGAAGCTGTGGCGCGCTGGCACCATGTGGAGCGGG 376
QY 286 GCGCGCGGCTTCTGGCTTCAACGCGGAGCGATCCGCTGCATGAGGCGCGGAGCATGCC 345
DB 377 GCGCGCGCTTCTGGCTTCAACGAGCGGCGCATGAGATAGCTTGTATGATACATACCA 436
QY 346 AAGTACGCTGCGCGGACCTGTGTACGCGGCGCGGCTACTGGGCGCAAGGCACTGCGCGAG 405
DB 437 AAGTACGCTGCGCGGACCTGTGTACGCGGCGCGGCTACTGGGCGCAAGGCACTGCGCGAG 496
QY 406 AACTGGCGCTGCGGAGCAACGCTGCTGCTTACTGGCGGAGCGGCGCGGCTGCTTTC 465
DB 497 AGTGTGCTGTACGAGGAGCAACGCTGCTGCTTCTGGGCTTACCGTTCATGGAGAGATCTTTC 555
QY 466 TACAGCGTGAACGAGGCGGAGCGGAGCTTCTCACTGCGGCGGCTGGGCGGCGCGG 525
DB 556 TACAGTATTTATGAGAGGAGGCAATATTGTTTCACTGTGGTAT-TAAGTCTCCGCTCTT 614
QY 526 CTCTGGGCGCTCATTTGATGCTTCAACGCAATCAACGAG-AGTGTGACGCTTCTGAGAGCGC 584
DB 615 CTCTGGGCGCTCATTTGATGCTTCAACGCAATCAACGAGAGTGAATTAATCTAGACGAGT 674
QY 585 CTCTGGCTGACACGCTGAGCGCGCGCGC 610
DB 675 GTTTGAGAGACCATGACCATGCGCC 700

RESULT 8
LOCUS A1579407 377 bp mRNA linear EST 05-APR-1999

DEFINITION UI-R-Ago-wy-b-01-0-UI.61 UI-R-Ago Rattus norvegicus cDNA clone
UI-R-Ago-wy-b-01-0-UI 3', mRNA sequence.

ACCESSION A1579407
VERSION A1579407.1 GI:4563783
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 377)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MBRRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.reagen.com) The following repetitive
elements were found in this cDNA sequence: 39-65,
>GC-rich low complexity
Seq primer: M13 Forward.
Location/Qualifiers
1..377
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Ago-wy-b-01-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-Ago"
/note="Vector: pUT73D-Pac (Pharmacia) with a modified
polylinker Site_1: Not 1; Site_2: Eco RI; The UI-R-Ago
library is a non-normalized library constructed from 13
dpc rat ventricle. The tag is a string of 6 nucleotides
present between the Not I site and the oligo-dT track.
The library was constructed as described by Bonaldo,
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Jim Lin, Department of Biology,
University of Iowa."

BASE COUNT 36 a 130 c 149 g 61 t 1 others
ORIGIN

Query Match 17.2%; Score 288.6; DB 9; Length 377;
Best Local Similarity 85.4%; Pred. No. 2.5e-40;
Matches 321; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 943 GCGCGGCGGCGCGCGGAGAGCTTCTGTGAGAGTGGGCGGCGGCGGCGGCGGCGGCGG 1002
DB 1 CGGCGGCGGCGCGCGGAGAGAGCTTGTGAGAGTGGGCGGCGGCGGCGGCGGCGGCGG 60
QY 1003 CCGCGGCGGCGGCGGCGGAGAGAGCTTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 1062
DB 61 CCGCGAGGTGAGGCTTGGGAGTCAAGCTGAGATCTTGGCGGCGGCGGCGGCGGCGGAG 120
QY 1063 CTGCGCGGCGGAG 1122
DB 121 CTGCGCGGCGGAGATCCGCTGCGCTGTGAGACGAGAGAGATCTGTGAGTGTGGCGGCGC 180
QY 1123 GGGCGCGGCGGAGGCGGAG 1182
DB 181 GGGCGCGGCGGAGGAG 240
QY 1183 CTCTGGGAGATCAACGAGAGAGTCCGCGCGGCGGCGGCTGTGTGTGTGTGTGTGTGTGT 1242
DB 241 CTGCGTGGGAG 300
QY 1243 CTGTGGGCGCTTCTTGTGCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1302
DB 301 CTGTGGGCGCTTCTTGTGCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 1303 CTGAGTCCAGGCGCTG 1318
DB 361 GTACAGTCCGATCTG 376

RESULT 9
LOCUS BB097995 593 bp mRNA linear EST 12-JUN-2000
DEFINITION UI-R-C4-anc-b-05-0-UI.61 UI-R-C4 Rattus norvegicus cDNA clone
UI-R-C4-anc-b-05-0-UI 3', mRNA sequence.

ACCESSION BE097995.1 GI:8488889
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS Donald, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 39-65,
 >GC rich#Low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 1..593
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C4-anc-b-05-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C4"
 /note="Vector: pTZ19-D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C4
 library is a subtracted library of a series, ultimately
 derived from a mixture of tissues from rat placenta,
 adult lung, brain, liver, kidney, heart, spleen, ovary,
 muscle, and 8, 12 and 18-day embryos. For a detailed
 description of the library from which this clone was
 derived, please visit our web site at
 ratseq.eng.uiowa.edu. This procedure has been previously
 described (Donald, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 TAG_SEQ=0
 BASE COUNT 94 a 162 c 200 g 136 t 1 others
 ORIGIN
 Query Match 16.64; Score 278.6; DB 10; Length 593;
 Best Local Similarity 86.34; Pred. No. 1.3e-38;
 Matches 308; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 943 CGCCGCGTGGCCCGGCGAGAGCTTCTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1002
 Db 1 CGGCGCGTGGCCCGGCGAGAGCTTGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 60
 QY 1003 CCGGCGCGTGGCCCGGCGAGAGCTTCTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1062
 Db 61 CCGGCGCGTGGCCCGGCGAGAGCTTCTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 120
 QY 1063 CTGCGCGCGAGCCGAGCGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1122
 Db 121 CTGCGCGCGAGTCCGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 180
 QY 1123 GGGCGCGTGTGTGAGCGCGGCGAGCGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1182

Db 181 GGGCGCGTGTGTGAGCGCGGCGAGCGAGCTTCTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 240
 QY 1183 CTGCGCGCGAGTCAACGCGGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1242
 Db 241 CTGCGCGCGAGTCAACGCGGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 300
 QY 1243 CTGCGCGCGTGTGTGAGCGCGGCGAGCGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 1299
 Db 301 CTGCGCGCGTGTGTGAGCGCGGCGAGCGCGTGTGTGAGGTGGCCGTCGCGGCGTGGCGGC 357

RESULT 10
 AM529983
 LOCUS
 DEFINITION UI-R-C4-alc-f-06-0-UI.81 UI-R-C4 Rattus norvegicus cDNA clone
 UI-R-C4-alc-f-06-0-UI 3', mRNA sequence.
 VERSION AM529983
 KEYWORDS
 SOURCE
 ORGANISM Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 392)
 AUTHORS Donald, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 39-65,
 >GC rich#Low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 1..392
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C4-alc-f-06-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C4"
 /note="Vector: pTZ19-D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C4
 library is a subtracted library of a series, ultimately
 derived from a mixture of tissues from rat placenta,
 adult lung, brain, liver, kidney, heart, spleen, ovary,
 muscle, and 8, 12 and 18-day embryos. For a detailed
 description of the library from which this clone was
 derived, please visit our web site at
 ratseq.eng.uiowa.edu. This procedure has been previously
 described (Donald, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=UI-R-C4
 TAG_TISSUE=corpus-strictum
 TAG_SEQ=CTAGG"
 TAG_SEQ=0
 BASE COUNT 38 a 132 c 156 g 66 t

Query Match 16.4%; Score 275.4; DB 9; Length 392;
 Best Local Similarity 85.7%; Pred. No. 4.7e-38;
 Matches 306; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 943 CGCCGCTGCGCCCGCGAGAGCTCTTCGTGAGGAGGCGCCGTCGCGCGCG 1002
 Db 1 CGGCGGCTGCGCCCGCGAGAGCTCTTCGTGAGGAGGCGCCGTCGCGCGCG 60

QY 1003 CGCGCGCGCTGCGCTTCGTGAGGAGGCGCCGTCGCGCGCG 1062
 Db 61 CGCGCGCGCTGCGCTTCGTGAGGAGGCGCCGTCGCGCGCG 120

QY 1063 CGCGCGCGCTGCGCTTCGTGAGGAGGCGCCGTCGCGCGCG 1122
 Db 121 CGCGCGCGCTGCGCTTCGTGAGGAGGCGCCGTCGCGCGCG 180

QY 1123 CGCGCGCGCTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCGCG 1182
 Db 181 CGCGCGCGCTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCGCG 240

QY 1183 CTCTGCGCGCTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCGCG 1242
 Db 241 CTCTGCGCGCTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCGCG 300

QY 1243 CTCTGCGCGCTTCGTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCG 1299
 Db 301 CTCTGCGCGCTTCGTGCGCGCGCGCGCGCGCGCTTCGTGAGGAGGCGCGCG 357

RESULT 11

BM440106

LOCUS

665 bp mRNA linear EST 01-FEB-2002
 pgrin.pk002.c6 Normalized Chicken Reproductive Tract cDNA library
 (pgrin) Gallus gallus cDNA clone pgrin.pk002.c6 5' similar to
 gb|AAK84420.1|AF400063_1 (AF400063) neutralized 1 [Mus musculus],
 mRNA sequence.

ACCESSION

BM440106

VERSION

BM440106.1

KEYWORDS

EST

SOURCE

Gallus gallus

ORGANISM

Gallus gallus (chicken)

REFERENCE

pgrin.pk002.c6

AUTHORS

Coburn, L.A. and Myers, Y.

TITLE

ESTs from Normalized Chicken Reproductive Tract cDNA library-

JOURNAL

University of Delaware and INRA, Tours-Poultry Unit Project

COMMENT

Unpublished

CONTACT

Contact: Larry A. Coburn

UNIVERSITY

University of Delaware

TOWNSEND

Townsend Hall, Newark, DE 19717, USA

TELEPHONE

Tel: 302-831-1335

FAX

Fax: 302-831-2822

EMAIL

Email: coburn@udel.edu, www.chickent.udel.edu.

LOCATION

Location/Qualifiers

FEATURES

1..665

SOURCE

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler and layer"

/db_xref="taxon:9031"

/clone="pgrin.pk002.c6"

/sex="Male and Female"

/tissue_type="Testis, ovary and oviduct"

/dev_stage="Various stages; embryonic, post-hatch, immature and sexually-mature"

/lab_host="E. coli EMDH10B"

/clone_lib="Normalized Chicken Reproductive Tract cDNA library (pgrin)"

/note="Vector: PCMVSPORT6; Library made from three total RNA pools from each tissue (testis 25%, ovary 25%, and oviduct 50% of final RNA pool); Single pass sequencing from 5'-end"

BASE COUNT 146 a 210 c 174 g 135 t
 ORIGIN

Query Match 15.8%; Score 264.4; DB 12; Length 665;
 Best Local Similarity 63.4%; Pred. No. 3.8e-36;
 Matches 425; Conservative 0; Mismatches 226; Indels 9; Gaps 1;

QY 557 CCGACGAGGTGACGCTTCGTGAGGAGGCGCCGTCGCGCGCG 616
 Db 5 CCGACGAGGTGACGCTTCGTGAGGAGGCGCCGTCGCGCGCG 64

QY 617 GCCAGGCGCGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCG 676
 Db 65 GCACTGCGCGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCG 124

QY 677 ACAACGAGCTGAGAACCAACCAAGGTGAGGCGCGCGCGCGCG 736
 Db 125 ACAATGAACTGAGAACCAACCAAGGTGAGGCGCGCGCGCGCG 184

QY 737 CCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 796
 Db 185 TGCCAGGACTGTGAGAACCAACCAAGGTGAGGCGCGCGCGCG 235

QY 797 GCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 856
 Db 236 GGCACGCTGTTCAGAGGAGGCGCGCGCGCGCGCGCGCGCGCG 295

QY 857 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 916
 Db 296 ATGACCGCGACATCATCTTTCTCAGAGCGAGAGGTGAGGCGCGCG 355

QY 917 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 976
 Db 356 GCAATGAGACTGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 415

QY 977 AGGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1036
 Db 416 AAGTGGAGACCTTTGGAGATGCGGATGCGGATGCGGATGCGGAT 475

QY 1037 ACCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
 Db 476 ATCCAAATGACTTTAAGCAAAAGAGTCCCAAGAGATCCGAGACTT 535

QY 1097 AAGAGTCTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1156
 Db 536 AAGAGTCTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 595

QY 1157 TCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1216
 Db 596 TCACAGTCTTGGCCAAACGCGCGAGGTGACATGAGGTGAGGAGG 655

QY 1217 TGCTGTGCGT 1226
 Db 656 TGAATGTGTGT 665

RESULT 12

BI404107

LOCUS

BI404107

DEFINITION

MI-P-CPI-nws-d-12-0-UI-31 MI-P-CPI Sus scrofa cDNA clone

ACCESSION

BI404107

VERSION

BI404107.1

KEYWORDS

EST

SOURCE

Sus scrofa (pig)

ORGANISM

Sus scrofa

REFERENCE

Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Euteleostomi; Cetartiodactyla; Suidae; Sus.

TITLE

Normalisation and subtraction: two approaches to facilitate gene

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 311)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
MEDLINE
PUBMED
8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab Clone distribution: Clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 39-65,
>GC rich/low complexity.
Seq primer: M13 Forward.
Location/Qualifiers
1. 311
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-G0-up-c-06-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_1lb="UI-R-G0"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-G0
library is a normalized library constructed from a
mixture of rat tissues (nodose ganglia, dorsal root
ganglia, and trigeminal ganglia). The tag is a string of
6 nucleotides present between the Not 1 site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996."

BASE COUNT 31 a 110 c 125 g 45 t
ORIGIN
Query Match 14.5%; Score 242.2; DB 9; Length 311;
Best Local Similarity 86.2%; Pred. No. 2.5e-32;
Matches 269; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

D 943 CGCCCGCTGGCGCCCGCGAGAGCTTCTGTGTGAGGTGGCGCTCCGGGCTGGCGCG 1002
1 CGGCGCTGGCGCCCGCGAGAGAGCTGTGTGTAATGGGGGCGCCGGGCTGGCGCGCG 60

D 1003 CCGCGCGCTGGCGCTTCCGATCATCTCTGTGAGTCTGCGCGCTGCGCGCATCCGAG 120
61 CCGCGAGCTGTGGCTTCCGATCATCTCTGTGAGTCTGCGCGCTGCGCGCATCCGAG 120

D 1063 CTGCGCGCGAGCCGAGCGGCTGTGACCGGAAAGTACTGGGCTGTGGCGCGCGCC 1122
121 CTGCGCGCGAGTCCGCTGGCGTGTGACCGGAAAGTACTGGGCTGTGGCGCGCGCC 180

D 1123 GGGCGCGTGGCGAGCGGCGGAGCGCGCTTCACTTCACTGCGCGCGCGCGAGCGT 1182
181 GGGCGCGTGGCGAGCGGCGGAGCGCGCTTCACTTCACTGCGCGCGCGCGAGCGT 240

D 1183 CTCTGGGAGTCAACGGGCGCTCGCGCGCGCGCTGTGTGCTGACACAGCGAGCGG 1242
241 CTCTGGGAGTCAACGGGCGCGCGCGCGCGCTGTGTGCTGACACAGCGAGCGG 300

D 1243 CTCTGGGAGCTT 1253
|||||

D 301 CTCTGGGAGCTT 311

RESULT 15
A2953956 362 bp DNA 11near GSS 27-APR-2001
LOCUS 2M0219K24F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M0219K24 F, genomic survey sequence.
ACCESSION A2953956
VERSION A2953956
KEYWORDS GI:13825183
SOURCE GSS.
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 362)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, B., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0219 Row: K Column: 24
Seq primer: CGTGTAAACGACGCGCGCT
Class: plasmid ends
High quality sequence stop: 362.
Location/Qualifiers
1. 362
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0219K24"
/sex="Female"
/lab_host="R. coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UUGC2M library"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gi|4732114|gb|AF12072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent R. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 43 a 136 c 119 g 64 t
ORIGIN
Query Match 13.1%; Score 220.2; DB 28; Length 362;
Best Local Similarity 80.8%; Pred. No. 1.6e-28;
Matches 269; Conservative 0; Mismatches 63; Indels 1; Gaps 1;

D 802 CGGCGCGCTGTGGCGCGGCGGCTACTGAGGCGGA-CTGCGCTTCCAGCAACAGCGG 860

